

The Effectiveness of Batterer Intervention Programs

A Literature Review & Recommendations for Next Steps

Patricia Cluss, PhD, and Alina Bodea, MD, MPH

©UNIVERSITY OF PITTSBURGH

March 2011

We thank the FISA Foundation for providing support for the preparation of this report.

PERMISSION GRANTED TO REPRODUCE AND DISSEMINATE
FOR EDUCATIONAL AND RESEARCH PURPOSES
EXCEPT FOR COPYRIGHTED ARTICLES IN APPENDIX E

TABLE OF CONTENTS

Section	Page
Background	2
Purpose of the Report	2
<i>Part 1: Review of the Literature</i>	
Review procedures	3
A note about scientific research design	4
Research formats and their ability to detect solid findings	4
Why it's hard to study Batterers Intervention Programs (BIPs)	6
<i>Part 2: Factors Involved in Understanding the Research Literature on BIPs</i>	
Types of interventions studied	8
Research sites	9
Populations studied	9
Outcome measures employed in BIP research	9
<i>Part 3: Major Findings in the Literature about BIP Effectiveness</i>	
Results of individual higher-quality studies	10
A note on the results of couples intervention studies	12
Results of meta-analyses and literature reviews	12
Other articles of interest	12
<i>Part 4: Conclusions and Recommendations for Next Steps</i>	
What we know so far	15
What we don't know yet	16
Recommendations	17
Conclusion	18
Appendix A: Measuring Intimate Partner Violence Victimization and Perpetration: A Compendium of Assessment Tools (partial report)	19
Appendix B: Reference list grouped by quality	56
Appendix C: Summaries high-quality articles	62
Appendix D: Summaries of medium-quality articles	79
Appendix E: Copies of high-quality articles	91

The Effectiveness of Batterer Intervention Programs

Background

Domestic violence (DV) is a major social and women's health concern. At least 85% of DV victims are women and approximately 1.5 million women in the U.S. experience physical or sexual violence from a current or former intimate partner each year. Nearly half of female victims sustain an injury at some point in the course of the abuse and 41% require medical care as a result of a physical assault by their partners. A long-term impact on health outcomes for women victims has also been documented, with reports of ongoing physical complaints such as gastrointestinal disorders, chronic pelvic pain, adverse pregnancy outcomes, decreased control over contraception and increased numbers of unintended pregnancies, among other chronic concerns. This health burden translates into higher health services use and costs. Female victims are twice as likely to use health care services than non-victims, with 2.5 times the health care costs. One study calculated costs of \$1.8 billion per year for direct medical care in the U.S. as a result of female victimization in relationships.

Much effort and expense in this country appropriately has gone into providing services and support for victims of domestic violence. However, focus on domestic violence batterers—the individuals who actually cause the problem—has lagged far behind the focus on victims in the areas of prevention and intervention. For example, a quick search of the comprehensive MEDLINE journal database in 2010 found 3606 articles on the topic of “domestic violence victims” and only 838 on the topic of “domestic violence batterers.” As a result of this lack of focus on batterers, relatively little is known about what constitutes a quality batterer's intervention program (BIP) model. Despite this lack of information, however, formal standards of care have been developed for BIPs and implemented in many states since the 1990s, based primarily on policy makers' beliefs about what constitutes a good program. Pennsylvania has been working on its own standards for many years, without reaching consensus about what they should contain.

Locally, it has been difficult to obtain reliable outcome data from existing local programs, so that organizations in the Pittsburgh area who work with batterers, such as the court system, and with DV victims, such as the local women's shelter programs, have been uncertain about the quality of BIPs. More information is needed both locally and nationally to inform the program development and evaluation of BIPs and to determine next steps for research into effective BIP programs.

Purpose of the Report

The purpose of this report is to provide a full and critical review of the effectiveness of batterers' intervention programs as evaluated in recent research published in peer-reviewed journals in the fields of medical and social science. We also make recommendations for next steps for researchers and BIP program developers, based on the results of the review. This review focuses only on group models of intervention, as there is very little research into individual models of batterer treatment.

Part 1: Review of the Literature

Review procedures:

A systematic search of MEDLINE and PsychINFO online databases was carried out during the months of July and August 2010 by the second author of this report. Several search combinations were employed using the keywords *batterers* and *perpetrators*. We “exploded” these terms in the search criteria, meaning that the search returned not only articles related to the selected keyword but also all of their more specific terms in the thematic areas of: partner abuse, domestic violence, intimate partner violence, intervention, treatment, and evaluation.

The search focused on literature published describing empirical studies published from 1990 to mid-2010 and resulted in a comprehensive list of relevant articles. Literature reviews and meta-analyses (see Format subsection below) published since 2000 were also identified. Other authoritative information sources were consulted for the purpose of identifying additional relevant publications. These sources included: the Centers for Disease Control's (CDC) *Center for Injury Prevention* webpage; the World Health Organization's *World Report on Violence and Health* (2002), and the *Mincava Electronic Clearinghouse* at the Minnesota Center Against Violence and Abuse. As the articles found through this initial search strategy were read and reviewed, additional relevant articles were identified from the reference sections of those articles.

Studies identified as a result of the search were then reviewed and categorized according to their format and quality.

Format: Articles on the effectiveness of BIPs were categorized as those reporting on single studies, systematic reviews of existing studies, or meta-analyses. Single studies are those where one intervention approach was evaluated either with or without a control group or where one approach was tested against another in a sample.

A systematic review is a literature review that tries to identify, evaluate, and integrate the research evidence relevant to a particular question or issue. So, our current report is a systematic literature review that includes results from other, earlier, systematic reviews that have been published in the literature.

A meta-analysis is similar to a systematic review in that it involves a thorough review of the literature, identifying studies on a specific research topic that are high quality. Additionally, a meta-analysis uses statistical methods to combine the results of the selected studies into one common measure called *effect size*. A meta-analysis has the advantage that it mathematically combines the results of many studies. Conclusions reached with a meta-analysis are based on a much larger total sample size than any of the individual studies included in the meta-analysis, thus increasing confidence in the accuracy of the results. Meta-analyses also specify the criteria that studies met to be included in the meta-analysis. In

this way, meta-analyses tend to make it clearer than with other kinds of formats the level of quality of the studies included.

Quality. studies were categorized according to the strength of the research design and the quality of outcomes into three groups: higher-quality studies; mid-quality studies; and lower-quality studies. Higher-quality and mid-quality studies are summarized in detail in this report and hard copies of high-quality articles are attached as appendices to the report. Additionally, a list of less well-designed studies and a list of studies relevant to the broader topic of BIPs, but not directed specifically at measuring effectiveness of an intervention approach are attached to the report as well.

As a result of the procedures described above, we are confident that this literature review provides an accurate representation of what is currently known about the effectiveness of BIPs. It reflects the critical themes of concern in the scientific community with respect to clinical, service and policy dimensions of BIPs and makes recommendations based on our findings for future directions for BIP research and program development.

A Note about Scientific Research Design

A premise inherent in this report is that, to serve as a useful model for broad-scale dissemination of best practices, a particular approach to a particular problem must have been subjected to an unbiased evaluation of outcomes. Such evaluations are best conducted according to scientific investigative techniques. One reason for adhering to this standard is that scientific techniques reduce as much as possible the inherent bias of the persons delivering the program. A second reason is that the use of scientific research design makes it more possible to specify the specific components of effective programs, so that others can model their programs after effective strategies.

Clinicians have a major role to play in this process by generating hypotheses about what kinds of treatments might be more effective than others based on their interactions with persons in the targeted population. Clinicians also have a major impact by delivering and disseminating best practice models, and by participating in the discussion about ways in which best practice models may need to be adapted for different subsets of the population.

Research formats and their ability to detect solid findings

A major focus of this review was to determine the scientific merit and, thereby, the soundness of findings for the studies published on BIPs' effectiveness. In order to investigate whether a certain BIP is effective in accomplishing its desired outcomes, the study itself has to be designed following rigorous formats. Generally, experimental designs provide the highest level of rigor. Experimental designs allow investigators to conclude with solid probability that the results observed are indeed due to key elements of the program being studied rather than to unrelated or irrelevant factors. Experimental designs can also provide evidence showing that there is not a strong probability that the targeted treatment is effective.

For an experimental design to be delivered in the most rigorous fashion, potential outside factors must be controlled in advance by the researchers. Researchers exert control in

several ways, but for our purposes here, the four most relevant are: homogeneity of the sample, use of control groups, random assignment to condition, and adequate sample size. Understanding these investigative methods will help the reader of this report make sense of the findings in the BIP literature.

Homogeneity of the sample: It is important for a study sample to contain participants who are similar to each other. So, for example, in a BIP effectiveness study, you might not want to have both court-mandated batterers and batterers who were voluntarily seeking treatment in your study. Most people would agree that the court-mandated type of batterer is probably different from the batterer who is voluntarily seeking treatment. If both types are included, a researcher may lose the power to find clear effects from the treatment because of factors related to who the participants are and not to the treatment itself.

Use of control groups: The purpose of research into interventions for domestic violence or anything else is to show cause and effect as clearly as possible. Investigators set up their studies so that a specific and well-defined treatment is delivered and so that any effects found are as likely as possible to be the result of the treatment--and not the result of any outside or unrelated factors. A key element in scientific design is the use of a control group. A control group is a group of people from the same population as the treatment group. The experimental group receives the treatment being studied. The control group receives no treatment. Then the investigator measures what happens to see whether the experimental group fares better in some predetermined ways than the control group. If so, the investigator can say with some confidence that the treatment is effective.

There has been a recent push for researchers to use “comparative effectiveness” research designs in situations in which it is not feasible or ethical to assign some participants to a no-treatment control group. This type of research design compares at least 2 (and sometimes more) treatments against each other, so that all participants receive some kind of treatment and the effects of the various treatments are measured scientifically. Comparative effectiveness studies can result in high quality results if done well. Sometimes, one of the treatments being studied can be considered a quasi-control group if the treatment components are relatively minimal and not expected to make a big difference with the population under study. An example of this kind of minimal treatment in BIP research would be a bibliotherapy or self-help treatment in which batterers were given printed information to read about the negative effects of battering.

Random assignment: One of the ways researchers control against the potential effect of unrelated factors on outcomes is by gathering a sample of persons in the population and then randomly assigning them to the various arms of the study (for example, to a treatment group vs. a control group or to treatment 1 vs. treatment 2). Random assignment means that the researchers themselves do not decide who gets which condition. Random assignment is particularly important because often the researchers implementing a study hope that the treatment they are studying will be found to be effective, or they are convinced ahead of time that it is. Without the standard of random assignment, they might want the most treatment-ready participants to be in the active treatment, for example, to show how well their treatment works under optimal conditions. Randomly assigning participants to a treatment and a control group makes it as likely as possible that those pre-existing biases on the part of the researcher do not affect the outcome of the study.

Adequate sample size: Consider this scenario: you have two participants volunteering for a BIP effectiveness study. You randomly assign one to the treatment condition and the other to the control condition. You provide treatment for the first participant and then check later to see how both participants are doing. Neither of the participants is doing too well. Both have been arrested again within weeks of the end of the study. Most reasonable people would conclude that, as disappointing as this result is, it does not really prove that the treatment under investigation is not effective. Why? Because most people understand that there may have been other factors besides the treatment that influenced the results for these two particular people. Perhaps the treatment is actually very effective, but Participant #1 was addicted to drugs in addition to being a batterer and couldn't focus on treatment. Perhaps Participant #2 was actually very motivated to improve his relationship with the victim and worked on his own (without treatment) to make things better.

It is obvious from this example that enrolling only two people in a study is not enough. How many participants, then, are enough to allow for confident interpretation of results? The answer, in general, is that more is better. Having 2000 participants in the study described above would be optimal. If the 1000 participants in the treatment do no better than the 1000 who did not receive treatment, then we would be able to conclude rather confidently that the treatment is not effective. Likewise, if the treated 1000 do show better outcomes than the control 1000, we can also be fairly confident that the treatment has shown effectiveness. However, most investigators cannot recruit 2000 participants for a study for practical and logistical reasons.

How do researchers know how many people are needed in a study in order to be confident of their eventual results? There is a statistical formula that results in a measure known as "power." Researchers can calculate "power" to figure out the minimum number of participants they would need to recruit in order to find a difference between treatment and control conditions if a true difference exists. Studies that have larger sample sizes and/or that have calculated the power statistic ahead of time are stronger studies than those that have not.

Why It's Hard to Study Batterers Intervention Programs

That said, one of the challenges of social science research is that some (or many) of the specific demands of experimental designs are difficult to achieve in real-life settings. In such settings (such as the batterer intervention program world), real-life needs and demands must take precedence over scientific method. A homogeneous sample, random assignment, adequate sample size and a control group often cannot be achieved for very practical reasons. For example, batterers who are court-mandated to an intervention program after an arrest for domestic violence cannot be assigned randomly to a no-treatment control group. All the batterers sent by the court must receive treatment. Instead of a sample size chosen ahead of time, investigators often have to make do with a "convenience sample"—whoever shows up for treatment in a given time frame. And samples in the real world may be made up of very different kinds of people, even though they share the same target problem. An example would be a treatment group that includes both court-mandated batterers and batterers who have not been involved in the legal system but who have sought out treatment on their own.

Design challenges: For this reason, many program evaluation studies make use of quasi-experimental or observational designs. “Quasi-experimental” means that some of the accepted research design components are met and some are not. Carefully designed quasi-experimental studies are able to control at least some of the potential structural and analytical shortcomings. Poorly designed quasi-experimental studies, though, introduce serious flaws that limit the validity of their findings.

An “observational study”, for our purposes, is one where the investigator observes an intervention or program and draws inferences about the possible effect of the treatment on participants. In an observational study, assignment of participants into a treated group versus a control group generally is outside the control of the investigator. A major challenge in conducting observational studies is to draw inferences that are acceptably free from influences of overt and hidden investigator biases. Results of purely observational studies are ones that inspire the least amount of confidence about their accuracy, compared to good-quality quasi-experimental and experimental studies.

Measurement challenges: The adequate identification and measurement of outcomes is an especially problematic dimension of BIP evaluation. How do we measure “He hasn’t changed” or “He’s doing better”? What to measure, when to measure it, and how to measure it are critical questions. Most commonly in BIP research to date, the primary outcome measures used are recidivism (rearrest) or reassault rates.

In order to obtain accurate recidivism rates, participants must be followed for a significant period of time after the end of a study, since only a small percentage of those who will go on to be rearrested for domestic assault do so within a few months of completing a study. However, it is difficult and expensive to find and follow up with batterers years later. Re-assault rates are also difficult to quantify, since not all assaults come to the attention of law enforcement (i.e. have a “paper trail” in the legal system) and both batterers’ and victims’ self-report of continued assaults may be inaccurate. Use of multiple measures of outcome is likely to provide better estimations of and a better understanding about critical change processes for batterers, but other methods are not generally in use in most of the literature to date. More information about measurement instruments is available in a compendium of assessment tools used for measuring domestic violence published by the Centers for Disease Control. (Thompson, Basile, Hertz & Sitterle, 2006) The batterer assessment section of the CDC publication is available in Appendix A of this report.

Other challenges: In BIP evaluation, other methodological and analytical difficulties exist in addition to those described above. These include: high drop out rates, the variability of program approaches, contents and jurisdictions; the multiple causes of domestic violence; and, probably, the existence of heterogeneous subtypes of persons who engage in domestic violence perpetration. According to one review (Eckhardt and colleagues, 2006), for example, between 40% to 60% of men mandated to BIP treatments either do not enroll in a group at all or drop out before completing a program. High drop out rates impede the researcher’s ability to describe outcomes adequately. Gathering of outcomes only from those who complete a program in which the norm is for participants to drop out is likely to bias results inaccurately in favor of program effectiveness.

Such difficulties add to the complexity of the task of adequately evaluating BIPs. One well-

known BIP investigator summarizes, “Evaluating the effectiveness of BIPs is a difficult and complex task that complicates the interpretation of evaluation results.” (Gondolf, 2004, p. 607)

Despite all these difficulties, it is generally considered that research on BIP effectiveness has been increasing in volume and quality over the past two decades. In order to continue this trend, there is growing consensus that standards of research on BIP effectiveness should consider: use of experimental or quasi-experimental designs with relevant controls; using broad definitions of abuse; use of multiple outcome measures, giving preference to victim reports over official reports of recidivism; completion of longer follow-up intervals for determining outcomes; and achievement of follow-up retention rates of at least 80%. Newer studies are also concerned about measuring the therapeutic integrity of programs as part of understanding the variables involved in their effectiveness. (ex: Saunders, 2009) These studies explicitly measure whether the treatment was delivered as it was supposed to be delivered during the investigation.

A note about the Duluth Model of Batterer Intervention: One particular model of batterer intervention, called the Duluth Model, is considered by many to be the standard for BIP programs. In fact, as of 2008, 45 states in the U.S. have legislated standards for BIPs and most of those mandate the use of at least some components of the feminist-psychoeducational Duluth Model as the treatment framework. The Duluth Model was developed by Minnesota Program Development, Inc., a nonprofit agency in Duluth, Minnesota. Their Domestic Abuse Intervention Project was the first multi-disciplinary program designed to address the issue of domestic violence. This program, conducted in 1981, coordinated the actions of a variety of agencies dealing with domestic conflict. The Duluth group developed the well-known Duluth Power and Control Wheel that makes use of concepts of institutionalized patriarchy to describe the power dynamics of batterer-victim relationships. Treatment in this model calls, among other components, for gender role resocialization--challenging batterer beliefs about men’s and women’s roles in society—and methods to reduce male dominance behaviors, as well as the prominence in treatment of victim safety. Many of the studies identified in this report investigated some version of a Duluth Model intervention.

Part 2: Factors Involved in Understanding the Research Literature on BIPs

➤ *Types of interventions studied*

Most studies we reviewed looked at judicially-mandated group interventions employing the feminist-psychoeducational (Duluth Model) or cognitive behavioral approaches or a combination of the two. These interventions typically were provided in all-male group formats and lasted anywhere from 12 to 52 weeks. One study (Morrel, Elliot, Murphy, Taft, 2003) compared a cognitive behavioral approach with a supportive group format. Another (Saunders, 1996) compared a Duluth Model approach with a process-psychodynamic intervention.

While most studies reviewed here compared common models of all-male group interventions,

a few compared results of couples intervention groups compared to men-only intervention groups (O'Leary, Heyman & Neidig, 1999; Dunford, 2000).

➤ *Research sites*

Most of the research described was conducted in single agency/single site locations and in contexts where the primary intention was to treat rather than study. A few multi-site studies exist. A recent state-wide evaluation of certified BIPs by MacLeod, Pi, Smith, and Rose-Goodwin (2009) examined whether variations between jurisdictions and BIPs predicted program outcome. See that article in Appendix C (#8) for a summary of their results. Gondolf (1999) studied BIPs in four U.S. cities and was able to discuss the comparability of treatment across sites, finding that 4 sites with moderate variability in treatment content found similar results across sites (see Appendix D #17).

➤ *Populations studied*

Most of the published rigorously-designed studies studied men who were referred for treatment by the courts. One study with a large sample size was conducted among Navy personnel (Dunford, 2000).

Despite the fact that a large proportion of BIP participants are racial and ethnic minorities (Gondolf, 2002), there are no effectiveness studies of culturally-tailored programs.

An interesting take on intervention is provided by the few qualitative studies that explored BIP participants' perception of interventions and the process of change (ex: Eckhardt, Holtzworth-Munroe, Norlander, Sibley and Cahill, 2008). These studies found generally low motivation of participants for intervention. Such findings may partly explain the issue of high drop out rates that undermine the ability of programs to describe outcomes adequately.

➤ *Outcome measures employed in BIP research*

In the literature reviewed, a program is considered effective if rates of aggressive behaviors are significantly reduced as a result of the intervention. Two types of sources are widely used to establish the reoccurrence of aggressive behaviors: official reports and victim reports. Official reports refer to either arrests for domestic violence or to official complaints made to the police. Research referenced by Feder and Wilson demonstrated that official reports capture only a small proportion of the abuse actually taking place (2005, p.252). Therefore, victim reports are considered a more accurate measure of aggression, but these may be affected by the victim's ongoing relationship with the batterer and other factors.

Several studies measured victims' reports of partners' abusive behavior using versions of the Conflict Tactics Scale (Strauss et al. 1996; see Appendix A). According to the meta-analysis of effect sizes by Feder and Wilson (2005), there are no statistically significant differences in outcome by type of report. However, the analysis by these authors suggested that victim reports add to the validity of studies and that quasi-experimental studies using victim reports probably show more adequate estimates of outcome than experimental studies using solely official records of reassault.

Other outcomes measured in the literature include: offenders and victims' attitudes about wife beating, about women, and responsibility; the likelihood of repeated abuse (Feder & Dugan, 2002); and standardized measures of aggression, global impression of change, communication behaviors, readiness to change, self-esteem and self-efficacy. (e.g., Morrel and colleagues, 2003). One study employed measures for degrees of violence, i.e., actual violence, violence threats, and terroristic threats (Edleson & Syers, 1990).

Part 3: Major Findings in the Literature about BIP Effectiveness

The research literature on BIP effectiveness includes a relatively small number of experimental and rigorous quasi-experimental studies and a relatively larger number of much less well-designed studies. The overarching observation in reviewing the literature is that **the more rigorous the methodology of evaluation studies, the less encouraging their findings.**

The results of the rigorous individual studies reviewed here, as well as most meta-analyses and systematic reviews conclude that **there is no solid empirical evidence for either the effectiveness or relative superiority of any of the current group interventions.** Across many rigorously conducted studies, treatment effects are small, if an effect exists at all, when comparing intervention to no intervention (control). Likewise, there is no significant, scientifically-verified difference between the effectiveness of different program models. There are intriguing results both about the possible positive effects of couples counseling interventions for selected subgroups of batterers and partners and also about the safety of victims who engage in couples intervention with an abusive partner.

Several examples of findings from high quality experimental outcome studies are summarized below. See Appendix B for a comprehensive list of high-, mid- and low-quality studies and articles, Appendix C for summaries of higher quality studies and articles, Appendix D for summaries of mid-quality studies, and Appendix E for copies of higher-quality articles.

Brief results of individual higher-quality studies

See Appendix C for details of randomization procedures, control groups and sample sizes for all studies described in this section.

Brannen and Rubin (1996) in a well-designed study, compared couples group intervention and gender-specific Duluth Model groups for men and victim support groups for women. See the article summary in Appendix C for a description of the elaborate safety system put in place for women participating in the study. These investigators found that couples intervention was more effective than gender-specific groups in reducing mild and severe physical abuse immediately after intervention by partner report and that it was particularly effective in couples where the batterer had a substance abuse problem. At 6 months, however, there was no difference in rearrest rates for batterers who had received couples vs. gender-specific intervention.

Dunford (2000) in another well-designed study with a military sample, compared a CBT men's group with a CBT couples groups and a rigorous monitoring (control) group. The study found no significant differences in effectiveness between couples intervention groups and men-only interventions.

Edleson and Syers (1990) compared a 12 and 32 session version of each of a structured educational model, a self-help group format, and a combination of the two, and found no significant differences between groups for the outcome measures of physical abuse and terroristic threats based on reports by partners.

Feder and Dugan (2002) randomly assigned 404 participants to a court-mandated program plus one year probation group or to a one year probation-only group and found no difference in subsequent violent behavior and rearrest at one year follow-up between the two groups. Additionally, men in the two groups showed no difference in attitudes or beliefs about domestic violence or in DV-related behavior.

MacLeod, Pi, Smith, and Rose-Goodwin (2009) conducted a state-wide evaluation of certified BIPs and examined whether variations between jurisdictions and BIPs predicted program outcome with over 1400 male offenders. They concluded that the strongest predictor of outcome was the individual characteristics of offenders rather than jurisdictional or BIP variations. In other words, the type of BIP treatment delivered did not exert a strong effect on outcome.

Morrel, Elliot, Murphy & Taft (2003), in a very well-designed study, compared a cognitive-behavioral group intervention against a support group intervention with 86 offenders and found no differences between the two interventions as measured by reports from partners at 6 months and official reports of recidivism at 2 and 3 years. Both groups were associated with significant decrease in physical, psychological and sexual abuse at follow up, calling into question the value of the more intensive CBT group compared to the relatively unstructured support group.

O'Leary, Heyman & Neidig (1999) compared a gender-specific individual intervention with a couples intervention. There were no significant differences in effectiveness between couples intervention groups and men-only interventions. Offenders reduced physical and psychological aggression significantly in all treatments. These investigators also measured victim fear and safety as part of participation in couples groups and found no added danger to victims who participated in such groups.

Saunders (1996) randomly assigned 213 participants to either a feminist-cognitive-behavioral model or process-psychodynamic groups and found no significant differences between the two types of interventions. The outcome measures were partners' reports of violence, fear, and relationship equality and recidivism. This study found that perpetrators with antisocial personalities had lower rates of recidivism in the feminist-cognitive therapy groups, while those with dependent personalities had reduced rates of recidivism in the process-psychodynamic groups.

Taylor, Davis & Maxwell (2001) conducted a randomized controlled trial comparing two

differing-length traditional group BIPs and a community service group with 376 male criminal court defendants assigned to a Duluth model intervention or a control group that completed 40 hours of community service. Results showed a significant reduction of violence according to police reports of recidivism measured at 6 and 12 months follow-up. However, there was no significant difference between the groups on partner reports of violence. There did appear to be an increase in the amount of “time to first official failure” (time interval before batterers were rearrested for DV assault) for offenders who attended the intervention group.

A note on the results of couples intervention studies

As Eckhardt and colleagues (2006) emphasize, the lack of difference found between couples therapy and either male-only CBT or a Duluth model therapy in the O’Leary and Dunford studies either means that neither of the interventions is particularly effective or that both couples therapy and male-only BIP group approaches are equally effective. The lack of a no-treatment control group with which to compare the two approaches keeps us from knowing which of these interpretations is correct, although the decreases in aggression reported in the O’Leary article are impressive.

IMPORTANT NOTE: Both the Dunford and O’Leary studies showing effectiveness for couples groups studied a very specific population of batterers that is not representative of other more commonly studied groups. In the O’Leary study, couples were a volunteer group (i.e. the men were not court-mandated) and were carefully screened to ensure that the physical injuries received by victims were not severe enough to need medical attention and that the victim was not afraid to be in a couples group with her abusive spouse. In the Dunford study, couples were recruited from a military population in which a strict structure and potential sanctions for batterers were firmly in place. The results from these studies cannot be generalized to other subgroups in the batterer population. The Brannen and Rubin study did use a court-referred but not court-mandated sample, since the victim in referred couples had to be willing to participate in order to be enrolled.

Results of meta-analyses and literature reviews

Several articles describe the results of meta-analytic investigations of good quality studies of BIP effectiveness and of comprehensive literature reviews. Summaries of these articles are available in Appendix C (meta-analyses: #1 and 7; literature reviews: # 4, 12, 13) and Appendix D (literature reviews #21 and 25). Findings from these informative articles are included in the Conclusions and Next Steps section below.

Other articles of interest

See the full references for these articles in Appendix B.

Ehrensaft and colleagues (2003): Intergenerational transmission of partner violence: A 20-year prospective study.

In a very strong study to try to identify the issues that a preventive approach to domestic violence would need to focus upon, these investigators followed a randomly selected and very large group of youth and their mothers for over 20 years and tested a developmental model of partner abuse, integrating the effects of witnessed family violence, child conduct problems and substance abuse. Originally recruited at ages 1 to 11, the 582 youth described in this article were now ages 17 to 28. Child conduct disorder was the strongest predictor of perpetrating partner violence as an adult, followed by witnessing DV as a child and receiving “power assertive punishment” as a child. Witnessing DV as a child was the strongest predictor for receiving DV as an adult. Investigators conclude that prevention efforts should focus on children with conduct disorder, those who witnessed DV in their home as children, and those who received excessive physical punishment as children. Their data support starting such prevention programs well before adolescence.

Gondolf (2009): Implementing mental health treatment for batterer program participants: Interagency breakdowns and underlying issues.

In this very interesting article, the investigator describes the barriers to a “community coordinated response” for batterers with mental health (MH) and addictions treatment needs. The article describes the results of an evaluation of a screening and referral system for BIP participants in the Pittsburgh area in which batterers receiving intervention via a court-ordered batterers program were screened for MH and alcohol problems and then referred to a MH clinic for follow up as part of their court-mandated treatment plan. In this large sample (N=1043), almost half screened positive for MH and/or alcohol problems. Problems were encountered in nearly every step of the implementation procedure: failures to screen per the established protocol at the BIP agency; inconsistent notification of results and referrals to BIP participants; lack of timely response by the MH clinic; insurance coverage difficulties; uncooperativeness of the batterers with the MH evaluation, resulting in lack of diagnosis that would substantiate the need for treatment; and significant problems with the courts, including judges’ inconsistent responses to referral noncompliance. The author identifies organizational and structural issues that contributed to these problems and makes recommendations for structural change and reorganization to improve a coordinated response for this population.

Hamberger, Lohr, Gottlieb (2000): Predictors of treatment dropout from a spouse abuse abatement program.

In another strong study, investigators first review the literature on what is known about predictors of BIP program attrition and then describe the results of a data analysis from a sample of 534 men enrolled in a BIP program, most of whom were court-mandated to treatment. They found that early drop out from programs (during assessment) was best predicted by high rates of previous police contact for violent crimes, failure to report an existing alcohol problem at intake, and paranoid personality characteristics. Late drop out (during treatment) was predicted by moderate/high rates of previous police contact for violent crimes and borderline personality characteristics. Interestingly, young violent offenders were more likely to complete treatment than others. The authors discuss the research and clinical implications of these results and suggest that batterers at risk for drop out can be identified at intake and adjustments can be made in program delivery to increase the likelihood that specific subtypes of batterers will complete treatment.

Holtzworth-Munroe, Meehan, Herron, Rezman, Stuart (2003): Do subtypes of martially violent men continue to differ over time?

The lead author in this paper previously has conducted and reported on the her research into subtypes of partner-violent men and has posited the idea that different forms or versions of treatment intervention may be needed for different subtypes. In this article, she and her colleagues examined whether men in the previously-identified subtypes continue to differ from each other over time--at 1.5 and 3-year follow up. The subtypes are: Family-Only (FO) batterers who are the least violent in the family compared to other subtypes, rarely violent outside the family, and show little psychopathology; 2) Borderline/Dysphoric (BD) batterers who engage in moderate to severe wife abuse, engage in some violence outside the family, and are the most psychologically distressed, including showing borderline personality characteristics; and 3) Generally violent/Antisocial (GVA) batterers who engage in moderate to severe family and extra-family violence and show evidence of other criminal behavior and/or substance abuse.

Investigators found that over the 3-year time period, BD and GVA men had the highest levels of reported partner violence and GVA men were least likely to have stopped being violent. FO men engaged in relatively low levels of marital violence and were the most likely to have stopped being violent in their relationships. These data are interesting because they potentially may help courts, interventionists and victims identify who is likely to continue being maritally violent. It also identifies a subset of batterers (FO group) whose partner violence does not tend to increase over time and who may, in fact, be able to discontinue violent and abusive behavior.

Rosenberg (2003): Voices from the group: Domestic violence offenders' experience of intervention.

This article describes the results of qualitative interviews with male and female DV offenders one year after completion of a 52-week court-mandated BIP. In general, program participants reported that relational factors in the group treatment (group support, alliance with the therapist) were most powerful in helping offenders reduce abusive behavior. Program provision of specific strategies for handling anger and other emotions and of positive interpersonal communication skills were also perceived as useful.

Smith and Randall (2007): Batterer intervention program: The victim's hope in ending the abuse and maintaining the relationship.

This article describes the results of qualitative interviews with female DV victims identifying their hopes and expectations for the results of participation by their violent/abusive partners in a BIP. Women described feeling an ongoing sense of oppression and injustice in their relationships, confusion about the best course of action to take, powerlessness, chronic fear, a sense of being trapped in the relationship, and strongly painful feelings about themselves and in general. Women tended to minimize the severity and meaning of their partners' abusive behaviors, engage in self-blame, maintain an emotional distance from the abuser and make unwanted life decisions, such as quitting their jobs, in response to the abuse. Once their fear reached a level that was no longer tolerable, victims called

the police, setting in motion a process that for the interviewed women resulted in court-mandated BIP treatment for their partners. Victims saw the BIP as the “last hope” for the relationship. They expressed their conviction that they would leave the batterer if abuse reoccurred and also expressed hope for change and faith that the batterer would change as a result of BIP intervention. Given the effectiveness data reviewed in this report and the high rates of drop out reported in court-mandated programs, victims’ hopes that a BIP program will “fix” their partners’ abusive behaviors appear to be unwarranted in the majority of cases. Clinical implications are discussed.

Taft and Murphy (2007): The working alliance in intervention for partner violence perpetrators: Recent research and theory.

These authors review the literature describing the effect of the “working alliance” between therapists and clients in batterer intervention work on program compliance and outcomes. Working alliance is defined as therapist and client agreement on the goals and tasks of therapy and the strength of the therapeutic bond. The authors conclude that the strength of the working alliance may be a significant factor in both compliance with BIP treatment and with treatment outcomes. They note that this work is in its very early stages and that the article provides only suggested directions for future research. The authors suggest more research into the effect of and possible alternative approaches to the use of confrontational behavior in batterer intervention, such as the use of motivational interviewing approaches that have been successful in substance abuse treatment and for other problems in which client resistance is typically high.

<p style="text-align: center;">Part 4: Conclusions and Recommendations for Next Steps</p>

There is a general consensus in the literature about what is known, what is not known, and what should be done next to improve the practice, policy, and research dedicated to BIPs.

What We Know So Far:

- There is **very little or no empirically demonstrated effectiveness of the widely available group interventions**, i.e., group programs for men, employing psycho-educational and/or cognitive behavioral approaches. Programs have at best very modest results.
- Intervention programs widely implemented by states and judicial systems that are based on **feminist-psychoeducational and/or cognitive-behavioral approaches lack empirical backing**.
- Perpetrators attending BIPs lack motivation for treatment.
- Mandated treatments seem **‘blind’ to the variability of needs and contexts** of participants.

- Theoretical approaches informing BIPs are based **less on empirical premises than on ideological positions**.

What We Don't Know Yet:

- An understanding of the complex etiology of domestic violence despite the abundance of theoretical models available.
- An adequate, empirically-supported understanding of how and why existing programs work or don't work.
- An understanding of the effectiveness of newer intervention approaches. There is growing research on such approaches as culturally tailored interventions, individually tailored interventions based on personality types, treatments for multiple etiologies, such as aggression in the context of substance abuse or mental conditions, yet these have not been properly evaluated to date.

Culturally-tailored interventions advocate the importance of social and cultural contexts in shaping attitudes to domestic violence, violent behaviors, and attitudes to treatment. As discussed by Whitacker and Niolon (2009, 182-183), there is inconclusive evidence on the differential effect of existing BIPs on culturally and racially differing men; there are a number of culturally-tailored programs available for African American, South-Western Asian, Native, and immigrant Latino men, yet these have not been rigorously tested for effectiveness.

Individually tailored interventions match psychological offender types to specific interventions. According to Whitacker and Niolon (2009, p.177-178) there are two most-cited typologies of domestic violence perpetrators; one of these is based on the frequency of violence and coercive control (Johnson, 1995); the other typology of abusive men was proposed by Holtzworth-Munroe and Stuart (2004) and is based on the frequency and generality of violence and on men's personality characteristics. Different types of violent men may benefit from different approaches, e.g., situational violence may respond better to couples counseling or anger management, while other intimate or patriarchal terrorism should be addressed by focusing on women's safety (Johnson, 1995, quoted by Whitaker & Niolon). In Saunders' (2006) evaluation of two interventions summarized in this review, it appeared that perpetrators with dependent personalities responded more favorably to the psychodynamic treatment, while those with antisocial traits responded better to the cognitive-behavioral model.

See also Eckhardt, Holtzworth-Munroe, Norlander, Sibley and Cahill's 2008 study on the relationship between readiness to change, perpetrator subtype, and treatment outcomes among men in treatment for assault; the authors found that BIP drop-out was higher for the borderline/dysphoric and generally violent/antisocial types; the same two types had also the highest re-arrest rates.

Motivational strategies are also a way of tailoring treatments to individual levels of

readiness, as postulated by the transtheoretical model of behavior change. See Eckhardt and colleagues (2008) for their findings on stages of change among men in treatment. See also studies by Taft and his collaborators (two are listed in the reference section) that demonstrate the value of motivational interviewing and the strength of the therapeutic alliance; a 2004 study by Taft, Murphy, Musser, and Remington entitled *Personality, interpersonal and motivational predictors of the working alliance in group cognitive-behavioral therapy for partner violent men*, found that motivational interviewing increased session attendance and reduced post-treatment intimate partner violence.

There is a solid research base documenting the relationship between domestic violence and substance abuse (see review by Whitaker and Niolon, 2009, p. 176-177), yet there are no studies of integrated models of intervention treating both violence and alcohol and/or substance abuse. A 1996 survey by Goldkamp, Welland, Collins, and White, *The role of drug and alcohol abuse in domestic violence and its treatment: Dade's County's domestic violence court experience* (quoted by Stuart, Temple and Moore, 2007), found that only 3% of men arrested for domestic violence were court mandated to also attend substance abuse treatment.

Despite the fact that couples intervention for domestic violence is prohibited in many states and is generally controversial, the effectiveness of couples therapy is supported by a number of studies, as reviewed by Stuart and colleagues (2007, p.562); these authors conclude that “for carefully selected clients, couples approaches may be helpful adjuncts to batterer intervention programs, may be beneficial subsequent to traditional batterer interventions, or in rare cases may be useful in lieu of batterer intervention” (2007, p. 562).

Recommendations

In light of these conclusions, a number of recommendations are unanimously formulated across concluding remarks of research studies as well as across reviews of the literature by such authors as Eckhardt and colleagues (2006), Stuart and colleagues (2007), Saunders (2008), Whitaker and Niolon (2009). Investigators should:

- Pilot a wide range of interventions, including couples interventions, and evaluate these carefully. Interventions should experiment with newer theoretical models and psychotherapeutic approaches and should have built-in research and evaluation components.
- Research should employ the most rigorous methodologies available, i.e., experimental designs with random assignment to intervention and control groups. Research should also be concerned with developing refined methodological instruments and procedures. Mixed method studies, combining quantitative and qualitative approaches and looking at programs at both state and local levels, should also be considered.
- Promising recent results from research on culturally tailored interventions, individually tailored treatment, substance abuse treatment, motivational strategies and couples

treatment suggest the value of investing more funding for program development and research in these areas.

- Developing service networks based on empirical evidence of effectiveness rather than on other motivations.
- Integrating BIPs into comprehensive integrated community services that can address adequately the multifaceted issue of domestic violence.

Conclusion

We conclude this report with a quote from Babcock and colleagues (2004), summarizing the results of their meta-analysis:

Because no one treatment model or modality has demonstrated superiority over the others, it is premature for states to issue mandates limiting the range of treatment options for batterers. Battering intervention agencies are more likely to improve their services by adding components or tailoring their treatments to specific clientele, than by rigidly adhering to any one curriculum in the absence of empirical evidence of its superior efficacy. Different types of batterers may preferentially benefit from specific forms of interventions, yet no controlled treatment-matching studies have been conducted to date. While a small number of studies have assessed group and couples' formats, no published studies to date have attempted to assess the efficacy of individual treatment for battering, although ... researchers are embarking on this frontier. (...) Promising directions for improving treatment efficacy include targeting treatments to specific subsamples, such as different ethnic minority groups, batterers who are chemically dependent, batterers at different motivational stages, different types of batterers (e.g., family-only, borderline, and antisocial/generally violent types), and women arrested for domestic violence. Treatment providers should develop alternative techniques and collaborate with researchers to evaluate their efficacy in an effort to develop evidence-based practice. To this end, researchers need to become an integral part of the coordinated community response to domestic violence.

Batterers' treatment is just one component of the coordinated community response to domestic violence. Police response, prosecution, probation, as well as treatment all affect recidivism of domestically violent partners. Even the best court-mandated treatment programs are likely to be ineffective in the absence of a strong legal response in initial sentencing and in sanctioning offenders who fail to comply with treatment. Even then, treatment may not be the best intervention for all batterers. Alternative sanctions should be developed and empirically tested along with alternative treatments (p.1048-1049).

APPENDIX A

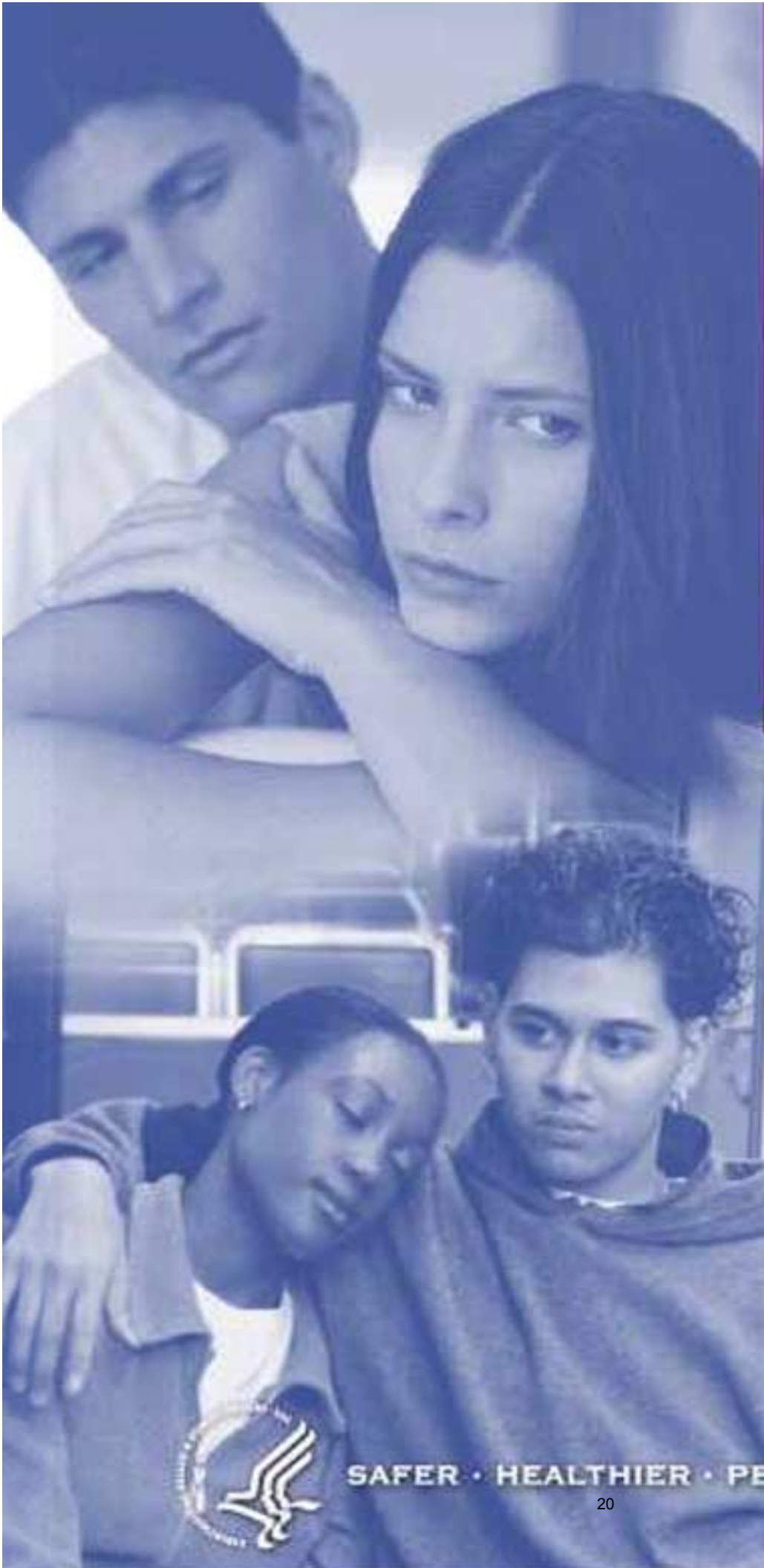
Measuring Intimate Partner Violence Victimization and Perpetration:

A Compendium of Assessment Tools

A Centers for Disease Control report
[pages i-iii, 1-3, 105-135]

Selected sections of the report are provided here on:
IPV Perpetration Assessment

The full report can be found at:
www.cdc.gov/ncipc/dvp/Compendium/IPV%20Compendium.pdf



Measuring Intimate Partner Violence Victimization and Perpetration:

A Compendium of Assessment Tools



SAFER • HEALTHIER • PEOPLE



Measuring Intimate Partner Violence Victimization and Perpetration: A Compendium of Assessment Tools is a publication of the National Center for Injury Prevention and Control of the Centers for Disease Control and Prevention.

Centers for Disease Control and Prevention

Julie L. Gerberding, MD, MPH, Director

National Center for Injury Prevention and Control

Ileana Arias, PhD, Director

Division of Violence Prevention

W. Rodney Hammond, PhD, Director



Suggested Citation: Thompson MP, Basile KC, Hertz MF, Sitterle D. Measuring Intimate Partner Violence Victimization and Perpetration: A Compendium of Assessment Tools. Atlanta (GA): Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, 2006.

Measuring Intimate Partner Violence Victimization and Perpetration:

A Compendium of Assessment Tools



Compiled and Edited by

Martie P. Thompson, PhD¹

Kathleen C. Basile, PhD²

Marci F. Hertz, MS²

Dylan Sitterle, BS¹

1 Department of Public Health Sciences
Clemson University

2 Department of Health and Human Services
Centers for Disease Control and Prevention
National Center for Injury Prevention and Control
Division of Violence Prevention

Atlanta, Georgia, 2006

Acknowledgments

We acknowledge the contributions of several CDC colleagues who reviewed the document and provided expert advice: Linda E. Saltzman, PhD; Thomas R. Simon, PhD; Linda L. Dahlberg, PhD; and James A. Mercy, PhD. We also extend thanks to Alida Knuth for editing, layout, and design assistance.

In addition, we acknowledge the following researchers who provided feedback on the scales identified for this compendium:

Ann Coker, PhD, University of Texas

Bonnie Fisher, PhD, University of Cincinnati

Mary Koss, PhD, University of Arizona

Jennifer Langhinrichsen-Rohling, PhD, University of South Alabama

Linda Marshall, PhD, University of North Texas

K. Daniel O’Leary, PhD, State University of New York at Stony Brook

Dan Saunders, PhD, University of Michigan

Jay Silverman, PhD, Harvard University

Paige Hall Smith, PhD, MPH, University of North Carolina at Greensboro

Dedication

We dedicate this compendium to the memory of Linda E. Saltzman, PhD, who strove in her professional work to improve the consistency of definitions and measurement of intimate partner violence.

Contents

Introduction	1
 Section A	
Physical Victimization Scales.....	5
 Section B	
Sexual Victimization Scales.....	27
 Section C	
Psychological/Emotional Victimization Scales	47
 Section D	
Stalking Victimization Scales.....	81
 Section E	
Physical Perpetration Scales	105
 Section F	
Sexual Perpetration Scales.....	115
 Section G	
Psychological/Emotional Perpetration Scales	123
 Section H	
Stalking Perpetration Scales.....	137
 Glossary	 151
 References	 152

Introduction

The Centers for Disease Control and Prevention (CDC) defines intimate partner violence (IPV) as actual or threatened physical, sexual, psychological, or stalking violence by current or former intimate partners (whether of the same or opposite sex). IPV is a major public health problem, reflected by both its prevalence and negative consequences. Researchers and prevention specialists are working to identify the factors that place intimate partners at risk for being victimized by or perpetrating violence, to find out which interventions are working, and to design more effective prevention programs.

National data suggest that IPV is perpetrated against both women and men, although most research indicates that women are more likely than men to be victimized by almost every type of IPV, including rape, physical assault, and stalking by an intimate partner (Tjaden and Thoennes 2000). The consequences of IPV are well documented and include substantial morbidity and mortality and physical and psychological health problems. Women are significantly more likely than men to be injured or killed by intimate partners. Approximately one in three females murdered in the United States is killed by a partner, whereas approximately one in twenty U.S. males murdered is killed by a partner (Puzone et al. 2000). Psychological consequences include posttraumatic stress disorder, depression, substance abuse, and suicidal behaviors (Caetano and Cunradi 2003; Campbell 2002; Coker et al. 2002; Hines and Malley Morrison 2001; Kaslow et al. 1998, 2002; Koss et al. 2003; Mechanic et al. 2000a.)

Purpose of the Compendium

This compendium provides researchers and prevention specialists with a compilation of tools designed to measure victimization from and perpetration of IPV. Many researchers are conducting studies to identify risk and protective factors for IPV and determine the consequences of victimization and perpetration. Others are working to design, implement, and evaluate interventions to reduce

IPV victimization and perpetration. The ability to accurately measure IPV is critical for the success of these research and intervention activities (Bachman 2000; Saltzman 2004).

In 1999, CDC published *Intimate Partner Violence Surveillance: Uniform Definitions and Recommended Data Elements* to improve and standardize data collected on IPV (Saltzman et al. 1999). Uniform and consistent definitions allow researchers and practitioners to assess the true prevalence of IPV, compare findings across studies, and determine the effectiveness of interventions. This compendium takes the next step by providing information on numerous scales for reliable and valid measurement of IPV.

Researchers and practitioners may find it challenging to identify which of the available scales are appropriate for measuring a particular type of IPV. This compendium provides professionals who are addressing this problem with easy access to a set of tools with demonstrated reliability and validity for measuring the self-reported incidence and prevalence of IPV victimization and perpetration. The compendium also identifies which scales are appropriate for measuring a given type of IPV.

What is Included in the Compendium?

Although this compendium includes more than 20 scales, it is not intended to be an exhaustive listing of available measures. The information is presented to help researchers and practitioners make informed decisions when choosing scales to use in their work. CDC does not endorse any particular scale presented in the compendium.

CDC used specific procedures to select scales for inclusion into the compendium. The process began with an intensive literature search and a review of articles published in violence-related and other journals over the past five years. This search identified a wide range of scales; CDC used several criteria to select a subset for inclusion in the compendium.

Scales had to be:

- published in a peer-reviewed journal or book,
- assessed for psychometric characteristics (with information on reliability, validity, or sensitivity available),
- created by the authors and not adapted from a preexisting scale,
- developed for research purposes,
- designed for direct participant response, and
- intended to assess actual violence rather than correlates, risk factors, or consequences of IPV.

If the original authors modified a scale and the modified version had published psychometric information, CDC included only the updated version. Those scales that were developed for screening or forensic purposes and scales completed by clinicians or through observational methods were excluded.

CDC consulted with a group of IPV research experts to review the instruments included in this compendium. To be as inclusive as possible of scales measuring all types of IPV, CDC selected experts who specialized in each of the four types of IPV (physical, sexual, psychological/emotional, and stalking).

This compendium includes a greater number of scales that assess victimization than those that assess perpetration. This likely reflects the field's historical focus on victimization. With the shift to research examining risk factors and evaluating perpetration interventions, it is likely that more scales assessing IPV perpetration will be forthcoming.

How is the Compendium Organized?

This compendium features scales measuring both victimization from and perpetration of IPV. Victimization scales are organized by physical violence victimization (Section A), sexual violence victimization (Section B), psychological/emotional abuse victimization (Section C), and stalking victimization (Section D). Perpetration scales are organized by physical violence perpetration (Section E), sexual violence perpetration (Section F),

psychological/emotional abuse perpetration (Section G), and stalking perpetration (Section H).

Each section begins with a table summarizing key information on each scale. The tables present information on the scale characteristics, target group or intended population, psychometric properties, authors, and year of publication.

For each included scale, the compendium provides scale items, response categories, scoring instructions, and the instructions provided to respondents at the beginning of the scale. Because all of the scales in this compendium have been previously published, CDC obtained permission to reprint each one from scale authors or publishing companies (when a scale was published in full in a journal). In some cases, publishers or authors required that CDC include a statement about a scale's copyright status. In those cases, this information is provided at the end of the scale. For two scales, publishers allowed only sample items to be reprinted. The full scales are available for purchase by contacting the publisher.

Some of the scales assess more than one type of violence. For example, a scale may assess both physical and psychological victimization. In these instances, the scale is repeated in the relevant sections, and the information on the target group and scale developer is the same. Psychometric data for each subscale are presented in the summary tables that open each section. To allow researchers to examine scale items for each type of violence in the context of the full scale, the complete scale is provided in each relevant category; item numbers pertaining to the relevant subscale are listed below each scale.

How to Use This Compendium

When selecting IPV scales for use, researchers should consider measurement issues such as how a particular scale operationally defines violence, how an intimate partner is defined, and what reporting time frame is used. The scales presented in this compendium assess different types of IPV. Some scales include items that assess only one type of violence, such as sexual violence or psychological abuse. Other scales are intended to assess more than one type of violence. Some scales assess both victimization from and perpetration of multiple forms of violence.

IPV scales also vary in terms of the population they are intended to assess. For example, some scales are limited to abused women, whereas other scales are intended for any woman with a current or former intimate partner. Some scales can be used to report on IPV in a current or former relationship, whereas other scales are intended for reporting on IPV perpetrated by former partners.

Intimate partner violence affects all racial and ethnic groups, and certain types of IPV may be more prevalent among African Americans, Hispanics, and Native American or Alaskan Natives (Tjaden and Thoennes 2000; Field and Caetano 2004). However, most scales in this compendium were not developed specifically for use with these or other minority populations. In most cases, reliability and validity information was obtained from largely non-hispanic white populations. For these reasons, the language used in most of the scales in this compendium may need to be adapted to be culturally or linguistically appropriate for some minority populations.

Some scales in this document are intended for use with adults; others are intended for use with adolescents or with any age group. The summary tables include specific information on intended age targets when that information is available.

None of the scales included in this compendium provide psychometric data specifically for same-sex couples. Researchers who wish to use the scales with same-sex couples should pilot test the scales with same-sex populations first.


The scales in this compendium also use a variety of reporting time frames. Researchers will need to decide which scales best suit their own research purposes. For example, if a researcher is interested in determining the prevalence of IPV among a specific population, then a scale that uses a lifetime reporting period may be most appropriate. If a researcher is interested in evaluating the effects of an intervention designed to reduce IPV victimization or perpetration, then the reporting time frame would need to coincide with the timing of the intervention.

Future Considerations

In the last two decades, IPV researchers have made great progress. However, several key areas need more attention. First, more research is needed to develop and test measures to assess perpetration of the various types of IPV, particularly sexual violence. Further, the field knows very little about the reliability and validity of the scales included here when used with different racial and ethnic populations and with same-sex relationships. It is CDC's hope that this document will encourage researchers to validate IPV victimization and perpetration measures in these understudied populations.

Section E

Physical Perpetration Scales



- E1. Abuse Within Intimate Relationships Scale
- E2. Abusive Behavior Inventory
- E3. Physical Abuse of Partner Scale
- E4. Revised Conflict Tactics Scales (CTS-2)
- E5. Safe Dates—Physical Violence Perpetration

Description of Measures

Perpetration Assessments					
Construct	Scale/Assessment	Characteristics*	Target Groups	Psychometrics	Developer
E. Physical Perpetration	E1. Abuse Within Intimate Relationships Scale (AIRS)	26-item scale that measures perpetration of psychological and physical abuse. There are 5 subscales: emotional abuse, deception, verbal abuse, overt violence, and restrictive violence.	Young adults.	Internal consistency: Overt violence = .86; Restrictive violence = .77.	Borjesson, Aarons, & Dunn, 2003 Copyright 2001
	E2. Abusive Behavior Inventory	30-item scale that measures the frequency of perpetration of physical and psychological abusive behaviors. The physical perpetration subscale includes 13 items (2 of which assess sexual abuse).	Male batterers.	Internal consistency: Physical abuse = .82. Evidence of criterion, convergent, and discriminant validity.	Shepard & Campbell, 1992 Copyright 1992
	E3. Physical Abuse of Partner Scale	25-item scale that measures the magnitude of physical abuse perpetrated against a spouse or partner.	Partners in dating, cohabiting, and marital relationships.	Internal consistency: > .90. Evidence of content and factorial validity.	Hudson, 1997 Copyright 1992
	E4. Revised Conflict Tactics Scales (CTS-2)	78-item scale that assesses both victimization and perpetration. The 39-item perpetration scale includes 5 subscales that measure physical assault, psychological aggression, sexual coercion, negotiation, and injury between partners. The physical assault subscale includes 12 items which can be grouped into 2 categories: minor and severe.	Partners in dating, cohabiting, and marital relationships.	Internal consistency: (men & women combined) Physical assault = .86. Evidence of convergent, discriminant and factorial validity.	Straus, Hamby, Boney-McCoy, & Sugarman, 1996; Straus, Hamby, & Warren, 2003 Copyright 2003
	E5. Safe Dates—Physical Violence Perpetration	16-item scale that measures physical perpetration in dating relationships.	Male and female students in grades 8-9	Internal consistency: 95.	Foshee, Linder, Bauman et al., 1996; Foshee et al., 1998

* Scale and subscale names in characteristics column are those that scale authors use and thus are not always consistent with CDC's terminology.

E1. Abuse within Intimate Relationships Scale (AIRS)

Please check the appropriate box for how often you have engaged in these behaviors.

Sample item of the 7 overt violence scale items:

I have physically attacked my partner.never once twice or more

☐ ☐ ☐

Sample item of the 3 restrictive violence scale items:

I have grabbed my partner's arm tightly.never once twice or more

☐ ☐ ☐

Copyright © 2001, Psychological Assessment Resources. Reproduced by special permission of the publisher, Psychological Assessment Resources (PAR), Inc., 16204 N. Florida Avenue, Lutz, FL 33549, from the Abuse Within Intimate Relationships Scale (AIRS) by Wiveca Borjesson, Gregory Aarons, and Michael Dunn. Further reproduction is prohibited without permission from PAR, Inc.

Scoring Instructions

To see entire scale, obtain permission to use, and obtain scoring information, contact:

Psychological Assessment Resources, Inc.
16204 N. Florida Avenue
Lutz, FL 33549

800-383-6595
813-968-3003
www.parinc.com

Reference

Borjesson WI, Aarons GA, Dunn ME. Development and confirmatory factor analysis of the Abuse Within Intimate Relationship Scale. *Journal of Interpersonal Violence* 2003;18:295–309.

E2. Abusive Behavior Inventory—Partner Form

Here is a list of behaviors that many women report have been used by their partners or former partners. We would like you to estimate how often you have used these behaviors during the past six months. Your answers are strictly confidential.

CIRCLE a number for each of the items listed below to show your closest estimate of how often it happened in your relationship with your partner or former partner during the past six months.

- 1 = Never
- 2 = Rarely
- 3 = Occasionally
- 4 = Frequently
- 5 = Very Frequently

- | | | | | | |
|---|---|---|---|---|---|
| 1. Called her names and/or criticized her | 1 | 2 | 3 | 4 | 5 |
| 2. Tried to keep her from doing something she wanted to do (example: going out with friends, going to meetings) | 1 | 2 | 3 | 4 | 5 |
| 3. Gave her angry stares or looks | 1 | 2 | 3 | 4 | 5 |
| 4. Prevented her from having money for her own use | 1 | 2 | 3 | 4 | 5 |
| 5. Ended a discussion with her and made the decision yourself | 1 | 2 | 3 | 4 | 5 |
| 6. Threatened to hit or throw something at her | 1 | 2 | 3 | 4 | 5 |
| 7. Pushed, grabbed, or shoved her | 1 | 2 | 3 | 4 | 5 |
| 8. Put down her family and friends | 1 | 2 | 3 | 4 | 5 |
| 9. Accused her of paying too much attention to someone or something else | 1 | 2 | 3 | 4 | 5 |
| 10. Put her on an allowance | 1 | 2 | 3 | 4 | 5 |
| 11. Used her children to threaten her (example: told her that she would lose custody, said you would leave town with the children) | 1 | 2 | 3 | 4 | 5 |
| 12. Became very upset with her because dinner, housework, or laundry was not ready when you wanted it or done the way you thought it should be | 1 | 2 | 3 | 4 | 5 |
| 13. Said things to scare her (examples: told her something “bad” would happen, threatened to commit suicide) | 1 | 2 | 3 | 4 | 5 |
| 14. Slapped, hit, or punched her | 1 | 2 | 3 | 4 | 5 |
| 15. Made her do something humiliating or degrading (example: begging for forgiveness, having to ask your permission to use the car or do something) | 1 | 2 | 3 | 4 | 5 |
| 16. Checked up on her (examples: listened to her phone calls, checked the mileage on her car, called her repeatedly at work) | 1 | 2 | 3 | 4 | 5 |

17. Drove recklessly when she was in the car	1	2	3	4	5
18. Pressured her to have sex in a way that she didn't like or want	1	2	3	4	5
19. Refused to do housework or childcare	1	2	3	4	5
20. Threatened her with a knife, gun, or other weapon	1	2	3	4	5
21. Spanked her	1	2	3	4	5
22. Told her that she was a bad parent	1	2	3	4	5
23. Stopped her or tried to stop her from going to work or school	1	2	3	4	5
24. Threw, hit, kicked, or smashed something	1	2	3	4	5
25. Kicked her	1	2	3	4	5
26. Physically forced her to have sex	1	2	3	4	5
27. Threw her around	1	2	3	4	5
28. Physically attacked the sexual parts of her body	1	2	3	4	5
29. Choked or strangled her	1	2	3	4	5
30. Used a knife, gun, or other weapon against her	1	2	3	4	5

Note: Item 21 was deleted from scale by its developers due to the low response rate and negative correlation with the total scale.

Copyright © 1992, Sage Publications, Thousand Oaks, CA. Used with permission.

Scoring Instructions

Physical abuse items include 6, 7, 14, 18, 20, 24, 25, 26, 27, 28, 29, and 30 (item 21 is not included in subscale computation). The mean score of these items is computed by summing the point values given in response to each item in the subscale and dividing by the applicable number of items. Higher scores are indicative of greater levels of physical abuse perpetration.

Reference

Shepard MF, Campbell JA. The Abusive Behavior Inventory: a measure of psychological and physical abuse. *Journal of Interpersonal Violence* 1992;7:291–305.

E3. Physical Abuse of Partner Scale (PAPS)

Name: _____ Today's Date: _____

This questionnaire is designed to measure the physical abuse you have delivered upon your partner. It is not a test, so there are no right or wrong answers. Answer each item as carefully and as accurately as you can by placing a number beside each one as follows.

- 1 = Never
- 2 = Very rarely
- 3 = A little of the time
- 4 = Some of the time
- 5 = A good part of the time
- 6 = Very frequently
- 7 = All of the time

- | | |
|--|---|
| _____ 1. I physically force my partner to have sex. | _____ 15. I knock my partner down and then kick or stomp him or her. |
| _____ 2. I push and shove my partner around violently. | _____ 16. I twist my partner's fingers, arms or legs. |
| _____ 3. I hit and punch my partner's arms and body. | _____ 17. I throw dangerous objects at my partner. |
| _____ 4. I threaten my partner with a weapon. | _____ 18. I bite or scratch my partner so badly that he or she bleeds or has bruises. |
| _____ 5. I beat my partner so hard he or she must seek medical help. | _____ 19. I violently pinch or twist my partner's skin. |
| _____ 6. I slap my partner around his or her face and head. | _____ 20. I hurt my partner while we are having sex. |
| _____ 7. I beat my partner when I'm drinking. | _____ 21. I hurt my partner's breast or genitals. |
| _____ 8. I make my partner afraid for his or her life. | _____ 22. I try to suffocate my partner with pillows, towels, or other objects. |
| _____ 9. I physically throw my partner around the room. | _____ 23. I poke or jab my partner with pointed objects. |
| _____ 10. I hit and punch my partner's face and head. | _____ 24. I have broken one or more of my partner's bones. |
| _____ 11. I beat my partner in the face so that he or she is ashamed to be seen in public. | _____ 25. I kick my partner's face and head. |
| _____ 12. I act like I would like to kill my partner. | |
| _____ 13. I threaten to cut or stab my partner with a knife or other sharp object. | |
| _____ 14. I try to choke or strangle my partner. | |

Copyright © 1992, James W. Garner and Walter W. Hudson, WALMYR Publishing Company. Illegal to photocopy or otherwise reproduce. Items from the PAPS provided by special permission of the publisher, WALMYR Publishing Company. Further reproduction is prohibited without permission from the publisher.

Scoring Instructions

To obtain permission to use and obtain scoring information, contact:

WALMYR Publishing Company
PO Box 12217
Tallahassee, FL 32317

(850) 383-0045
walmyr@walmyr.com.

Reference:

Hudson WW. The WALMYR assessment scales scoring manual. Tallahassee (FL): WALMYR Publishing Company; 1997.

E4. Revised Conflict Tactics Scales (CTS-2)

No matter how well a couple gets along, there are times when they disagree, get annoyed with the other person, want different things from each other, or just have spats or fights because they are in a bad mood, are tired, or for some other reason. Couples also have many different ways of trying to settle their differences. This is a list of things that might happen when you have differences. Please circle how many times you did each of these things in the past year, and how many times your partner did them in the past year. If you or your partner did not do one of these things in the past year, but it happened before that, circle "7."

How often did this happen?

1 = Once in the past year

2 = Twice in the past year

3 = 3–5 times in the past year

4 = 6–10 times in the past year

5 = 11–20 times in the past year

6 = More than 20 times in the past year

7 = Not in the past year, but it did happen before

0 = This has never happened

Sample of 2 of the 12 physical assault scale items:

I pushed or shoved my partner. 1 2 3 4 5 6 7 0

I punched or hit my partner with something that could hurt. 1 2 3 4 5 6 7 0

Copyright © 2003, Western Psychological Services. Reproduced by special permission of the publisher, Western Psychological Services, 12031 Wilshire Boulevard, Los Angeles, CA 90025, from the Revised Conflict Tactics Scales (CTS-2) by Murray Straus, Sherry Hamby, Sue Boney-McCoy, and David Sugarman. Further reproduction is prohibited without permission from Western Psychological Services.

Scoring Instructions

To see entire scale, obtain permission to use, and obtain scoring information, contact:

Western Psychological Services
Attn. Susan Weinberg
12031 Wilshire Boulevard
Los Angeles, CA 90025

weinberg@wpspublish.com
(800) 648-8857

References

Straus MA, Hamby SL, Boney-McCoy S, Sugarman DB. The Revised Conflict Tactics Scale (CTS2): development and preliminary psychometric data. *Journal of Family Issues* 1996;17:283–316.

Straus MA, Hamby SL, Warren WL. The Conflict Tactics Scale handbook. Los Angeles (CA): Western Psychological Services; 2003.

E5. Safe Dates—Physical Violence Perpetration

How many times have you ever done the following things to a person that you have been on a date with? Only include when you did it to him/her first. In other words, don't count it if you did it in self-defense. Please circle one number on each line.

	10 or more times	4 to 9 times	1 to 3 times	Never
1. Scratched them	3	2	1	0
2. Slapped them	3	2	1	0
3. Physically twisted their arm	3	2	1	0
4. Slammed or held them against a wall	3	2	1	0
5. Kicked them	3	2	1	0
6. Bent their fingers	3	2	1	0
7. Bit them	3	2	1	0
8. Tried to choke them	3	2	1	0
9. Pushed, grabbed, or shoved them	3	2	1	0
10. Dumped them out of a car	3	2	1	0
11. Threw something at them that hit them	3	2	1	0
12. Burned them	3	2	1	0
13. Hit them with my fist	3	2	1	0
14. Hit them with something hard besides my fist	3	2	1	0
15. Beat them up	3	2	1	0
16. Assaulted them with a knife or gun	3	2	1	0

Scoring Instructions

Point values are indicated above. The physical perpetration scale score is calculated by summing the point values of the 16 responses. The mean value can also be obtained by dividing the summed responses by the number of items (16). Higher scores are indicative of greater physical perpetration. Scores can also be categorized such that "0" = no physical perpetration, "1" = 1 to 3 times, and "2" indicates perpetration 3 or more times.

References

Foshee VA, Bauman KE, Arriaga XB, Helms RW, Koch GG, Linder GF. An evaluation of Safe Dates, an adolescent dating violence program. *American Journal of Public Health* 1998;88:45–50.

Foshee VA, Linder GF, Bauman KE, et al. The Safe Dates project: theoretical basis, evaluation design, and selected baseline findings. *American Journal of Preventive Medicine* 1996;12:39–47.

Section F

Sexual Perpetration Scales



F1. Revised Conflict Tactics Scales (CTS-2)

F2. Sexual Experiences Survey (SES)

Description of Measures

Perpetration Assessments					
Construct	Scale/Assessment	Characteristics*	Target Groups	Psychometrics	Developer
F. Sexual Perpetration	F1. Revised Conflict Tactics Scales (CTS-2)	78-item scale that assesses both victimization and perpetration. The 39-item perpetration scale includes 5 subscales that measure physical assault, psychological aggression, sexual coercion, negotiation, and injury between partners. The sexual coercion subscale includes 7 items that can be grouped into minor and severe categories based on whether or not physical force is used to achieve coercion.	Partners in dating, cohabiting, and marital relationships.	Internal consistency (men & women combined): Sexual coercion = .87.	Straus, Hamby, Boney-McCoy, & Sugarman, 1996; Straus, Hamby, & Warren, 2003 Copyright 2003
	F2. Sexual Experiences Survey (SES)—Perpetration Version	10-item scale that measures 4 types of sexual perpetration.	Male college students.	Internal consistency: Males = .89. Test-retest correlation = .93. Evidence of criterion validity. At the time of this publication, the SES was undergoing revision, but new psychometric data were yet to be published. Contact Mary Koss for updates at: mpk@email.arizona.edu	Koss & Gidycz, 1985; Koss, Gidycz, & Wisniewski, 1987; Koss & Oros, 1982

* Scale and subscale names in characteristics column are those that scale authors use and thus are not always consistent with CDC's terminology.

F1. Revised Conflict Tactics Scales (CTS-2)

No matter how well a couple gets along, there are times when they disagree, get annoyed with the other person, want different things from each other, or just have spats or fights because they are in a bad mood, are tired, or for some other reason. Couples also have many different ways of trying to settle their differences. This is a list of things that might happen when you have differences. Please circle how many times you did each of these things in the past year, and how many times your partner did them in the past year. If you or your partner did not do one of these things in the past year, but it happened before that, circle “7.”

How often did this happen?

1 = Once in the past year

5 = 11–20 times in the past year

2 = Twice in the past year

6 = More than 20 times in the past year

3 = 3–5 times in the past year

7 = Not in the past year, but it did happen before

4 = 6–10 times in the past year

0 = This has never happened

Sample of 2 of the 7 sexual coercion scale items:

I used force (like hitting, holding down, or using a weapon) to make my partner have oral or anal sex.	1	2	3	4	5	6	7	0
--	---	---	---	---	---	---	---	---

I insisted on sex when my partner did not want to (but did not use physical force).	1	2	3	4	5	6	7	0
---	---	---	---	---	---	---	---	---

Copyright © 2003, Western Psychological Services. Reproduced by special permission of the publisher, Western Psychological Services, 12031 Wilshire Boulevard, Los Angeles, CA 90025, from the Revised Conflict Tactics Scales (CTS-2) by Murray Straus, Sherry Hamby, Sue Boney-McCoy, and David Sugarman. Further reproduction is prohibited without permission from Western Psychological Services.

Scoring Instructions

To see entire scale, obtain permission to use, and obtain scoring information, contact:

Western Psychological Services
Attn. Susan Weinberg
12031 Wilshire Boulevard
Los Angeles, CA 90025

weinberg@wpspublish.com
800) 648-8857

References

Straus MA, Hamby SL, Boney-McCoy S, Sugarman DB. The Revised Conflict Tactics Scale (CTS2): development and preliminary psychometric data. *Journal of Family Issues* 1996;17:283–316.

Straus MA, Hamby SL, Warren WL. *The Conflict Tactics Scale handbook*. Los Angeles (CA): Western Psychological Services; 2003.

F2. Sexual Experiences Survey (SES)—Perpetration Version

On the following pages are questions about your sexual experiences from age 14 on.

1. Have you engaged in sex play (fondling, kissing, or petting, but not intercourse) when she didn't want to because you overwhelmed her with continual arguments and pressure? ☐ Yes ☐ No
 If No, continue with question 2.
 If Yes:
 1a. About how many times has it happened (from age 14 on)? 1 2 3 4 5 or more
 1b. How many times last school year (September to September)? 0 1 2 3 4 5 or more

2. Have you engaged in sex play (fondling, kissing, or petting but not intercourse) when she didn't want to because you used your position of authority (boss, teacher, camp counselor, supervisor) to make her? ☐ Yes ☐ No
 If No, continue with question 3.
 If Yes:
 2a. About how many times has it happened (from age 14 on)? 1 2 3 4 5 or more
 2b. How many times last school year (September to September)? 0 1 2 3 4 5 or more

3. Have you engaged in sex play (fondling, kissing, or petting but not intercourse) when she didn't want to because you threatened or used some degree of physical force (twisting her arm, holding her down, etc.) to make her? ☐ Yes ☐ No
 If No, continue with question 4.
 If Yes:
 3a. About how many times has it happened (from age 14 on)? 1 2 3 4 5 or more
 3b. How many times last school year (September to September)? 0 1 2 3 4 5 or more

4. Have you attempted sexual intercourse (get on top of her, attempt to insert your penis) when she didn't want to by threatening or using some degree of force (twisting her arm, holding her down, etc.), but intercourse did not occur? ☐ Yes ☐ No
 If No, continue with question 5.
 If Yes:
 4a. About how many times has it happened (from age 14 on)? 1 2 3 4 5 or more
 4b. How many times last school year (September to September)? 0 1 2 3 4 5 or more

5. Have you attempted sexual intercourse (get on top of her, attempt to insert your penis) when she didn't want to by giving her alcohol or drugs, but intercourse did not occur? ☐ Yes ☐ No
- If No, continue with question 6.
- If Yes:
- 5a. About how many times has it happened (from age 14 on)? 1 2 3 4 5 or more
- 5b. How many times last school year (September to September)? 0 1 2 3 4 5 or more
6. Have you engaged in sexual intercourse when she didn't want to because you overwhelmed her with continual arguments and pressure? ☐ Yes ☐ No
- If No, continue with question 7.
- If Yes:
- 6a. About how many times has it happened (from age 14 on)? 1 2 3 4 5 or more
- 6b. How many times last school year (September to September)? 0 1 2 3 4 5 or more
7. Have you engaged in sexual intercourse when she didn't want to because you used your position or authority (boss, teacher, camp counselor, supervisor) to make her? ☐ Yes ☐ No
- If No, continue with question 8.
- If Yes:
- 7a. About how many times has it happened (from age 14 on)? 1 2 3 4 5 or more
- 7b. How many times last school year (September to September)? 0 1 2 3 4 5 or more
8. Have you engaged in sexual intercourse when she didn't want to because you gave her alcohol or drugs? ☐ Yes ☐ No
- If No, continue with question 9.
- If Yes:
- 8a. About how many times has it happened (from age 14 on)? 1 2 3 4 5 or more
- 8b. How many times last school year (September to September)? 0 1 2 3 4 5 or more
9. Have you engaged in sexual intercourse when she didn't want to because you threatened or used some degree of physical force (twisting her arm, holding her down, etc.) to make her? ☐ Yes ☐ No
- If No, continue with question 10.
- If Yes:
- 9a. About how many times has it happened (from age 14 on)? 1 2 3 4 5 or more
- 9b. How many times last school year (September to September)? 0 1 2 3 4 5 or more

10. Have you engaged in sex acts (anal or oral intercourse or penetration by objects other than the penis) when she didn't want to because you threatened her or used some degree of physical force (twisting her arm, holding her down, etc.) to make her? ☐ Yes ☐ No

If No, continue with question 11.

If Yes:

10a. About how many times has it happened (from age 14 on)? 1 2 3 4 5 or more

10b. How many times last school year (September to September)? 0 1 2 3 4 5 or more

11. Did you answer "Yes" to any of the questions 1-10? ☐ Yes ☐ No

If Yes:

11a. Look back to the questions 1-10 in this section. What is the highest question number to which you marked "Yes"?

1 2 3 4 5 6 7 8 9 10

Note: Scale can be used to assess sexual violence perpetrated by non-intimates.

Scoring Instructions

Respondents are classified according to the most severe sexual perpetration that they reported, ranging from no sexual victimization to rape. Men are classified as perpetrators of rape if they answered "yes" to items 8, 9, or 10. Men are classified as perpetrators of sexual coercion if they answered "yes" to items 6 or 7 but not to any higher numbered items. Men are classified as perpetrators of attempted rape if they answered "yes" to items 4 or 5 but not to any higher numbered items. Men are classified as perpetrators of sexual contact if they answered "yes" to Items 1, 2, or 3 but not to any higher numbered items.

References


Koss MP, Gidycz CA. Sexual Experience Survey: reliability and validity. *Journal of Consulting and Clinical Psychology* 1985;53:422–423.

Koss MP, Gidycz CA, Wisniewski N. The scope of rape: incidence and prevalence of sexual aggression and victimization in a national sample of higher education students. *Journal of Consulting and Clinical Psychology* 1987;55:162–170.

Koss MP, Oros CJ. Sexual Experience Survey: a research instrument investigating sexual aggression and victimization. *Journal of Consulting and Clinical Psychology* 1982;50:455–457.

Section G

Psychological/ Emotional Perpetration Scales



- G1. Abuse Within Intimate Relationships Scale
- G2. Abusive Behavior Inventory
- G3. Multidimensional Measure of Emotional Abuse
- G4. Non-Physical Abuse of Partner Scale (NPAPS)
- G5. Revised Conflict Tactics Scales (CTS-2)
- G6. Safe Dates—Psychological Abuse Perpetration

Description of Measures

Perpetration Assessments					
Construct	Scale/Assessment	Characteristics*	Target Groups	Psychometrics	Developer
G. Psychological/Emotional Perpetration	G1. Abuse Within Intimate Relationships Scale (AIRS)	26-item scale that measures perpetration of psychological and physical abuse. There are 5 subscales: emotional abuse, deception, verbal abuse, overt violence, and restrictive violence.	Young adults.	Internal consistency: Emotional abuse = .87; Deception = .80; Verbal abuse = .73.	Borjesson, Aarons, & Dunn, 2003 Copyright 2001
	G2. Abusive Behavior Inventory	30-item scale that measures the frequency of physical and psychological abusive behaviors. The psychological perpetration subscale includes 17 items.	Male batterers.	Internal consistency: Psychological abuse = .79 to .88. Evidence of convergent, discriminant, and criterion validity.	Shepard & Campbell, 1992 Copyright 1992
	G3. Multidimensional Measure of Emotional Abuse	28-item scale (reduced from 54 items) that measures restrictive engulfment, hostile withdrawal, denigration, and dominance/intimidation.	College students reporting on current or past dating relationships.	Internal consistency: Restrictive engulfment = .84; Hostile withdrawal = .88; Denigration = .89; Dominance/Intimidation = .83. Evidence of convergent and discriminant validity.	Murphy & Hoover, 1999; Murphy, Hoover, & Taft, 1999
	G4. Non-Physical Abuse of Partner Scale (NPAPS)	25-item scale that measures the magnitude of perceived non-physical abuse inflicted on a spouse or partner.	Partners in dating, cohabiting, and marital relationships.	Internal consistency: > .90. Evidence of content and factorial validity.	Hudson, 1997 Copyright 1992
	G5. Revised Conflict Tactics Scales (CTS-2)	78-item scale that assesses both victimization and perpetration. The 39-item perpetration scale includes 5 subscales that measure physical assault, psychological aggression, sexual coercion, negotiation, and injury between partners. The psychological aggression subscale includes 8 items that assess verbal and symbolic acts that are intended to cause fear or psychological distress.	Partners in dating, cohabiting, and marital relationships.	Internal consistency (men & women combined): Psychological aggression = .79.	Straus, Hamby, Boney-McCoy, & Sugarman, 1996; Straus, Hamby, & Warren, 2003 Copyright 2003
	G6. Safe Dates—Psychological Abuse Perpetration	14-item scale that measures psychological perpetration in dating relationships.	Male and female students in grades 8-9.	Internal consistency: .95.	Foshee, Linder, Bauman et al., 1996; Foshee et al., 1998

* Scale and subscale names in characteristics column are those that scale authors use and thus are not always consistent with CDC's terminology.

G1. Abuse within Intimate Relationships Scale (AIRS)

Please check the appropriate box for how often you have engaged in these behaviors.

Sample item of the 7 emotional abuse scale items:

I have purposely insulted my partner.never once twice or more
☐ ☐ ☐

Sample item of the 4 deception scale items:

I have kept secrets from my partner.never once twice or more
☐ ☐ ☐

Sample item of the 5 verbal abuse scale items:

I have ignored my partner.never once twice or more
☐ ☐ ☐

Copyright © 2001, Psychological Assessment Resources. Reproduced by special permission of the publisher, Psychological Assessment Resources (PAR), Inc., 16204 N. Florida Avenue, Lutz, FL 33549, from the Abuse Within Intimate Relationships Scale (AIRS) by Wiveca Borjesson, Gregory Aarons, and Michael Dunn. Further reproduction is prohibited without permission from PAR, Inc.

Scoring Instructions

To see entire scale, obtain permission to use, and obtain scoring information, contact:

Psychological Assessment Resources, Inc.
 16204 N. Florida Avenue
 Lutz, FL 33549

800-383-6595
 813-968-3003
www.parinc.com

Reference

Borjesson WI, Aarons GA, Dunn ME. Development and confirmatory factor analysis of the Abuse Within Intimate Relationship Scale. *Journal of Interpersonal Violence* 2003;18:295–309.

G2. Abusive Behavior Inventory—Partner Form

Here is a list of behaviors that many women report have been used by their partners or former partners. We would like you to estimate how often you have used these behaviors during the past six months. Your answers are strictly confidential.

CIRCLE a number for each of the items listed below to show your closest estimate of how often it happened in your relationship with your partner or former partner during the past six months.

- 1 = Never
- 2 = Rarely
- 3 = Occasionally
- 4 = Frequently
- 5 = Very Frequently

- | | | | | | |
|---|---|---|---|---|---|
| 1. Called her names and/or criticized her | 1 | 2 | 3 | 4 | 5 |
| 2. Tried to keep her from doing something she wanted to do (example: going out with friends, going to meetings) | 1 | 2 | 3 | 4 | 5 |
| 3. Gave her angry stares or looks | 1 | 2 | 3 | 4 | 5 |
| 4. Prevented her from having money for her own use | 1 | 2 | 3 | 4 | 5 |
| 5. Ended a discussion with her and made the decision yourself | 1 | 2 | 3 | 4 | 5 |
| 6. Threatened to hit or throw something at her | 1 | 2 | 3 | 4 | 5 |
| 7. Pushed, grabbed, or shoved her | 1 | 2 | 3 | 4 | 5 |
| 8. Put down her family and friends | 1 | 2 | 3 | 4 | 5 |
| 9. Accused her of paying too much attention to someone or something else | 1 | 2 | 3 | 4 | 5 |
| 10. Put her on an allowance | 1 | 2 | 3 | 4 | 5 |
| 11. Used her children to threaten her (example: told her that she would lose custody, said you would leave town with the children) | 1 | 2 | 3 | 4 | 5 |
| 12. Became very upset with her because dinner, housework, or laundry was not ready when you wanted it or done the way you thought it should be | 1 | 2 | 3 | 4 | 5 |
| 13. Said things to scare her (examples: told her something “bad” would happen, threatened to commit suicide) | 1 | 2 | 3 | 4 | 5 |
| 14. Slapped, hit, or punched her | 1 | 2 | 3 | 4 | 5 |
| 15. Made her do something humiliating or degrading (example: begging for forgiveness, having to ask your permission to use the car or do something) | 1 | 2 | 3 | 4 | 5 |
| 16. Checked up on her (examples: listened to her phone calls, checked the mileage on her car, called her repeatedly at work) | 1 | 2 | 3 | 4 | 5 |

17. Drove recklessly when she was in the car	1	2	3	4	5
18. Pressured her to have sex in a way that she didn't like or want	1	2	3	4	5
19. Refused to do housework or childcare	1	2	3	4	5
20. Threatened her with a knife, gun, or other weapon	1	2	3	4	5
21. Spanked her	1	2	3	4	5
22. Told her that she was a bad parent	1	2	3	4	5
23. Stopped her or tried to stop her from going to work or school	1	2	3	4	5
24. Threw, hit, kicked, or smashed something	1	2	3	4	5
25. Kicked her	1	2	3	4	5
26. Physically forced her to have sex	1	2	3	4	5
27. Threw her around	1	2	3	4	5
28. Physically attacked the sexual parts of her body	1	2	3	4	5
29. Choked or strangled her	1	2	3	4	5
30. Used a knife, gun, or other weapon against her	1	2	3	4	5

Note: Item 21 was deleted from scale by its developers due to the low response rate and negative correlation with the total scale.

Copyright © 1992, Sage Publications, Thousand Oaks, CA. Used with permission.

Scoring Instructions

Psychological abuse items include 1, 2, 3, 4, 5, 8, 9, 10, 11, 12, 13, 15, 16, 17, 19, 22, and 23. The mean score of these items is computed by summing the point values for the items and dividing by the applicable number of items. Higher scores are indicative of greater psychological abuse perpetration.

Reference

Shepard ME, Campbell JA. The Abusive Behavior Inventory: a measure of psychological and physical abuse. *Journal of Interpersonal Violence* 1992;7:291–305.

G3. Multidimensional Measure of Emotional Abuse

The following questions ask about the relationship with your partner or ex-partner. Please report how often each of these things has happened in the last six months. Please circle a number using the scale below to indicate how often you have done each of the following things, and a number to indicate how often your partner has done each of the following things. Indicate how many times you have done this where it says “you”, and how many times your partner has done this where it says “your partner”. If you or your partner did not do one of these things in the past 6 months, but it has happened before that, circle “7”.

1 = Once 4 = 6-10 times 7 = Never in the past six months, but it has happened before
 2 = Twice 5 = 11-20 times 0 = This has never happened
 3 = 3-5 times 6 = More than 20 times

	Once	Twice	3-5 times	6-10 times	11-20 times	More than 20 times	Never in the past six months, but it has happened before	This has never happened
1. Asked the other person where they had been or who they were with in a suspicious manner								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
2. Secretly searched through the other person's belongings								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
3. Tried to stop the other person from seeing certain friends or family members								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
4. Complained that the other person spends too much time with friends								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
5. Got angry because the other person went somewhere without telling him/her								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
6. Tried to make the other person feel guilty for not spending enough time together								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0

	Once	Twice	3-5 times	6-10 times	11-20 times	More than 20 times	Never in the past six months, but it has happened before	This has never happened
7. Checked up on the other person by asking friends or relatives where they were or who they were with								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
8. Said or implied that the other person was stupid								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
9. Called the other person worthless								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
10. Called the other person ugly								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
11. Criticized the other person's appearance								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
12. Called the other person a loser, failure, or similar term								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
13. Belittled the other person in front of other people								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
14. Said that someone else would be better partner (better spouse, better girlfriend or boyfriend)								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
15. Became so angry that they were unable or unwilling to talk								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
16. Acted cold or distant when angry								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0

	Once	Twice	3-5 times	6-10 times	11-20 times	More than 20 times	Never in the past six months, but it has happened before	This has never happened
17. Refused to have any discussion of a problem								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
18. Changed the subject on purpose when the other person was trying to discuss a problem								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
19. Refused to acknowledge a problem that the other person felt was important								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
20. Sulked or refused to talk about an issue								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
21. Intentionally avoided the other person during a conflict or disagreement								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
22. Became angry enough to frighten the other person								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
23. Put his/her face right in front of the other person's face to make a point more forcefully								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
24. Threatened to hit the other person								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
25. Threatened to throw something at the other person								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
26. Threw, smashed, hit, or kicked something in front of the other person								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0

	Once	Twice	3-5 times	6-10 times	11-20 times	More than 20 times	Never in the past six months, but it has happened before	This has never happened
27. Drove recklessly to frighten the other person								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0
28. Stood or hovered over the other person during a conflict or disagreement								
You	1	2	3	4	5	6	7	0
Your Partner	1	2	3	4	5	6	7	0

Scoring Instructions

Items can be used to create one total scale score and four subscale scores. The 7-item Restrictive Engulfment subscale consists of items 1-7. The 7-item Denigration subscale consists of items 8-14. The 7-item Hostile Withdrawal subscale consists of items 15-21. The 7-item Dominance/Intimidation subscale consists of items 22-28. Higher scores are indicative of greater levels of emotional abuse.

Two types of scores for the total scale score and for the subscale scores can be computed. One scoring method involves assigning a score of 0 if the respondent reports never having done a particular act, and a score of 1 if a respondent reports having done a particular act. A second scoring method involves using frequency counts in specific intervals of time. In this scoring method, a score of 7 is recoded to 0, and then the 0-6 point items are summed.

References

Murphy, CM, Hoover, SA. Measuring emotional abuse in dating relationships as a multifactorial construct. *Violence and Victims* 1999;14: 39-53.

Murphy, CM, Hoover, S, Taft, C. *The Multidimensional Measure of Emotional Abuse: Factor structure and subscale validity*. Toronto: Association for the Advancement of Behavior Therapy; 1999.

G4. Non-Physical Abuse of Partner Scale (NPAPS)

Name: _____ Today's Date: _____

This questionnaire is designed to measure the non-physical abuse you have delivered upon your partner. It is not a test, so there are no right or wrong answers. Answer each item as carefully and as accurately as you can by placing a number beside each as follows.

- 1 = Never
- 2 = Very rarely
- 3 = A little of the time
- 4 = Some of the time
- 5 = A good part of the time
- 6 = Very frequently
- 7 = All of the time

- | | |
|---|---|
| _____ 1. I make fun of my partner's ability to do things. | _____ 14. I demand that my partner stay home. |
| _____ 2. I expect my partner to obey. | _____ 15. I don't want my partner to work or go to school. |
| _____ 3. I become very upset and angry if my partner says that I have been drinking too much. | _____ 16. I don't want my partner socializing with his or her female friends. |
| _____ 4. I demand my partner to perform sex acts that he or she does not enjoy or like. | _____ 17. I demand sex whether my partner wants it or not. |
| _____ 5. I become very upset if my partner's work is not done when I think it should be. | _____ 18. I scream and yell at my partner. |
| _____ 6. I don't want my partner to have any male friends. | _____ 19. I shout and scream at my partner when I'm drinking. |
| _____ 7. I tell my partner he or she is ugly and unattractive. | _____ 20. I order my partner around. |
| _____ 8. I tell my partner to hop to it when I give him or her an order. | _____ 21. I have no respect for my partner's feelings. |
| _____ 9. I expect my partner to hop to it when I give him or her an order. | _____ 22. I act like a bully towards my partner. |
| _____ 10. I insult or shame my partner in front of others. | _____ 23. I frighten my partner. |
| _____ 11. I become angry if my partner disagrees with my point of view. | _____ 24. I treat my partner like he or she is a dimwit. |
| _____ 12. I carefully control the money I give my partner. | _____ 25. I'm rude to my partner. |
| _____ 13. I tell my partner that he or she is dumb or stupid. | |

Copyright © 1992, James W. Garner and Walter W. Hudson, WALMYR Publishing Company. Illegal to photocopy or otherwise reproduce. Items from the NPAPS provided by special permission of the publisher, WALMYR Publishing Company. Further reproduction is prohibited without permission from the publisher.

Scoring Instructions

To obtain permission to use and obtain scoring information, contact:

WALMYR Publishing Company
PO Box 12217
Tallahassee, FL 32317

(850) 383-0045
walmyr@walmyr.com.

Reference

Hudson WW. The WALMYR assessment scales scoring manual. Tallahassee (FL): WALMYR Publishing Company; 1997.

G5. Revised Conflict Tactics Scales (CTS-2)

No matter how well a couple gets along, there are times when they disagree, get annoyed with the other person, want different things from each other, or just have spats or fights because they are in a bad mood, are tired, or for some other reason. Couples also have many different ways of trying to settle their differences. This is a list of things that might happen when you have differences. Please circle how many times you did each of these things in the past year, and how many times your partner did them in the past year. If you or your partner did not do one of these things in the past year, but it happened before that, circle “7.”

How often did this happen?

- 1 = Once in the past year
- 2 = Twice in the past year
- 3 = 3–5 times in the past year
- 4 = 6–10 times in the past year
- 5 = 11–20 times in the past year
- 6 = More than 20 times in the past year
- 7 = Not in the past year, but it did happen before
- 0 = This has never happened

Sample of 2 of the 8 psychological aggression scale items:

I called my partner fat or ugly.	1	2	3	4	5	6	7	0
I shouted or yelled at my partner.	1	2	3	4	5	6	7	0

Copyright © 2003, Western Psychological Services. Reproduced by special permission of the publisher, Western Psychological Services, 12031 Wilshire Boulevard, Los Angeles, CA 90025, from the Revised Conflict Tactics Scales (CTS-2) by Murray Straus, Sherry Hamby, Sue Boney-McCoy, and David Sugarman. Further reproduction is prohibited without permission from Western Psychological Services.

Scoring Instructions

To see entire scale, obtain permission to use, and obtain scoring information, contact:

Western Psychological Services
Attn. Susan Weinberg
12031 Wilshire Boulevard
Los Angeles, CA 90025

weinberg@wpspublish.com
(800) 648-8857

References

Straus MA, Hamby SL, Boney-McCoy S, Sugarman DB. The Revised Conflict Tactics Scale (CTS-2): development and preliminary psychometric data. *Journal of Family Issues* 1996;17:283–316.

Straus MA, Hamby SL, Warren WL. *The Conflict Tactics Scale handbook*. Los Angeles (CA): Western Psychological Services; 2003.

G6. Safe Dates—Psychological Abuse Perpetration

How often have you done the following things to someone you have ever had a date with? Please circle one number on each line.

	Very often	Sometimes	Seldom	Never
1. Damaged something that belonged to them.....	3	2	1	0
2. Said things to hurt their feelings on purpose.	3	2	1	0
3. Insulted them in front of others.	3	2	1	0
4. Threw something at them that missed.	3	2	1	0
5. Would not let them do things with other people.	3	2	1	0
6. Threatened to start dating someone else.	3	2	1	0
7. Told them they could not talk to someone of the opposite sex.	3	2	1	0
8. Started to hit them but stopped.	3	2	1	0
9. Did something just to make them jealous.	3	2	1	0
10. Blamed them for bad things I did.	3	2	1	0
11. Threatened to hurt them.	3	2	1	0
12. Made them describe where they were every minute of the day.	3	2	1	0
13. Brought up something from the past to hurt them.	3	2	1	0
14. Put down their looks.	3	2	1	0

Scoring Instructions

The psychological abuse perpetration score is calculated by summing responses across all 14 items. Summed scores are recoded as follows:

- 0 = 0 and indicates no perpetration.
- 1–5 = 1 and indicates mild psychological abuse.
- 6–9 = 2 and indicates moderate psychological abuse.
- 10 and greater = 3 and indicates severe psychological abuse.

References

Foshee VA, Bauman KE, Arriaga XB, Helms RW, Koch GG, Linder GF. An evaluation of Safe Dates, an adolescent dating violence program. *American Journal of Public Health* 1998;88:45–50.

Foshee VA, Linder GF, Bauman KE, et al. The Safe Dates project: theoretical basis, evaluation design, and selected baseline findings. *American Journal of Preventive Medicine* 1996;12:39–47.

APPENDIX B

Quick Reference Guide: Articles Grouped by Quality

APPENDIX B

Quick Reference Guide: Articles Grouped by Quality

Higher-Quality Articles

Summaries of these studies are compiled in Appendix C.

1. Babcock, J.C., Green, C.E, Robie, C. (2004). Does batterers' treatment work? A meta-analytic review of domestic violence treatment. *Clinical Psychology Review*, 23(8), 1023-1053.
[A meta-analysis that reviews 22 experimental and quasi-experimental studies; total sample size from all studies=3857]
2. Brannen, S.J., Rubin, A. (1996). Comparing the effectiveness of gender-specific and couples groups in a court-mandated spouse abuse treatment program. *Research on Social Work Practice*, 6, 405-424.
[Random assignment+; control group--no, but this is a comparative effectiveness study; sample size=49 couples]
3. Dunford, F.W. (2000). The San Diego Navy experiment: an assessment of intervention for men who assault their wives. *Journal of Consulting and Clinical Psychology*, 68(3), 468-476.
[Random assignment+; control group+; sample size=861 married couples]
4. Eckhardt, C.I., Murphy, C., Black, D., Suhr, L. (2006). Intervention programs for perpetrators of intimate partner violence; Conclusions from a clinical research perspective. *Public Health Reports*, 121, 389-381.
[A literature review of 7 experimental studies and several previous reviews of batterer intervention; details regarding sample size of studies is not indicated in the published report]
5. Edleson, J.L., & Syers, M. (1990). Relative effectiveness of group treatments for men who batter. *Social Work Research and Abstracts*, 26(2), 10-18.
[Random assignment+; control group--no, but this is a comparative effectiveness study; the minimal self-help treatment condition might be considered a control group; sample size=283]
6. Feder, L., Dugan, L., (2002). A test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward experiment. *Justice Quarterly*, 19(2), 343-375.
[Random assignment+; control group+; sample size=404]
7. Feder, L., Wilson, D.B. (2005). A meta-analytic review of court-mandated batterer intervention programs: Can courts affect abusers' behavior? *Journal of Experimental Criminology*, 1, 239-262.
[A meta-analysis that reviews 4 experimental and 6 quasi-experimental studies; total combined sample size is not noted in the publication]

8. MacLeod, D., Pi, R., Smith, D., Rose-Goodwin, L. (2009). Batterer intervention systems in California. An evaluation. Judicial Council of California, Office of the Courts. Full text available online at www.courtinfo.ca.gov/reference/documents/batterer-report.pdf.
[A state program evaluation that reviews BIPs in 5 judicial jurisdictions in California; total sample size =~1400]
9. Morrel, T.M., Elliot, J.D., Murphy, C.M., Taft, C.T. (2003). Cognitive Behavioral and Supportive Group treatments for partner-violent men. *Behavior Therapy*, 34, 77-95.
[Random assignment+ (for quasi-random assignment); control group--no, but this is a comparative effectiveness study; the supportive therapy treatment condition might be considered a control group; sample size=86]
10. O'Leary, D.K., Heyman, R.E., Neidig, P.H. (1999). Treatment of wife abuse: A comparison of gender-specific and conjoint approaches. *Behavior Therapy*, 30, 475-505.
[Random assignment+; control group--no, but this is a comparative effectiveness study; sample size=75 male-female couples]
11. Saunders, D.G. (1996). Feminist-cognitive-behavioral and process-psychodynamic treatments for men who batter: Interactions of abuser traits and treatment models. *Violence and Victims*, 11(4), 393-414.
[Random assignment+; control group--no, but this is a comparative effectiveness study; sample size=218]
12. Stover, C.S., Meadows, A.M., Kaufman, J. (2009). Interventions for intimate partner violence: review and implications for evidence-based practice. *Professional Psychology: Research and Practice*, 40(3), 223-233.
[A literature review of 11 experimental studies of batterer intervention; total combined sample size=2358 for treatment program participants (excludes one study with 4032 participants who received a non-counseling court intervention)]
13. Stuart, G.L, Temple, J.R, Moore, T.M. (2007). Improving batterer intervention programs through theory-based research. *JAMA*, 298(5), 560-562.
[A succinct mini-literature review published in the prestigious Journal of the American Medical Association that makes a number of empirically-driven recommendations for improving programs and policy-makers ability to make informed decisions about effective treatments]
14. Taylor, B.G., Davis, R.C., Maxwell, C.D. (2001). The effects of a group batterer treatment program: A randomized experiment in Brooklyn. *Justice Quarterly*, 18(1), 171-201.
[Random assignment+; control group+; sample size=376]

Mid-Quality Articles

Summaries of these articles are compiled in Appendix D.

15. Dutton, D.G., Bodnarchuk, M., Kropp, R., Hart, S.D., Ogloff, J.R.P. (1997). Wife assault treatment and criminal recidivism: An 11-year follow-up. *International Journal of Offender Therapy and Comparative Criminology*, 41, 9-23.
[Random assignment-; control group-; sample size=446]
16. Eckhardt, C., Holtzworth-Munroe, A., Norlander, B., Sibley, A., Cahill, M. (2008). Readiness to change, partner violence subtypes, and treatment outcomes among men in treatment for partner assault. *Violence and Victims*, 23(4), 446-475.
[Random assignment-; control group-; sample size=199]
17. Gondolf, E.W. (1999). A comparison of four batterer intervention Systems. Do court referral, program length, and services matter? *Journal of Interpersonal Violence*, 14(1), 41-61.
[Random assignment-; control group-; sample size=840]
18. Gondolf, E.W. (2000). A 30-month follow-up of court-referred batterers in four cities. *International Journal of Offender Therapy and Comparative Criminology*, 44(1), 111-128.
[Random assignment-; control group-; sample size=618]
19. Gondolf, E.W. (2004). Evaluating batterer counseling programs: A difficult task showing some effects and implications. *Aggression and Violent Behavior*, 9, 605-631.
[Random assignment-; control group-; sample size=840 batterers and their female partners]
20. Gordon, J.A., Moriarty, L.J. (2003). The effects of domestic violence batterer treatment on domestic violence recidivism. The Chesterfield County experience. *Criminal Justice and Behavior*, 30(1), 118-134.
[Random assignment-; control group+ (but it is a non-equivalent control group); sample size=248]
21. Saunders, D.G. (2008). Group interventions for men who batter: a summary of program descriptions and research. *Violence and Victims*, 23(2), 156-172.
[A literature review of batterer intervention studies ; details regarding sample size of studies is not indicated in the published report]
22. Snow-Jones, A., D'Agostino, R.B., Jr., Gondolf, E.W., Heckert, A. (2004). Assessing the effect of batterer program completion on reassault using propensity scores. *Journal of Interpersonal Violence*, 19(9), 1002-1020.
[Random assignment-; control group-; sample size=633]
23. Snow-Jones A., Gondolf, E.W. (2001). Time-varying risk factors for reassault among batterer program participants. *Journal of Family Violence*, 16(4), 345-359.
[Random assignment-; control group-; sample size=308]

24. Taft, C.T., Murphy, C.M., King, D.W., Musser, P.H., DeDeyn, J.M. (2003). Process and treatment adherence factors in group cognitive-behavioral therapy for partner violent men. *Journal of Consulting and Clinical Psychology*, 71(4), 812-820.
[Random assignment-; control group-; sample size=107]
25. Whitaker, D. J. & Niolon, P.H. (2009). Advancing Interventions for Perpetrators of Physical Partner Violence: Batterer Intervention Programs and Beyond. In D. J. Whitaker and J. R. Lutzker, *Preventing partner violence: Research and evidence-based intervention strategies*. Washington, DC: American Psychological Association, pp. 169-192.
[A comprehensive literature review of batterer intervention approaches and studies; lack of tables with easy-to-read summaries of details and common components of studies makes it somewhat less useful than other reviews.]

Lower-Quality Articles

Articles are listed here for reference purposes but the full articles are not included in the report.

26. Coulter, M., VandeWeerd, C. (2009). Reducing domestic violence and other criminal recidivism: effectiveness of a multilevel batterers intervention program. *Violence and Victims*, 24(2), 139-152.
[RA-; control group-; sample size=17,999; problem: compares outcomes for program completers with outcomes for program drop outs and inappropriately infers the difference in results was due to the effect of treatment rather than to other possible factors]
27. Gondolf, E.W. (2009). Outcomes from referring batterer program participants to mental health treatment. *Journal of Family Violence*, 24, 577-588.
[Random assignment-; control group-; sample size=148; problem: the high rates of noncompliance with mental health referral compromise the quality of results]
28. Tutty, L.M., Bidgood, B.A., Rothery, M.A., Bidgood, P. (2001). An evaluation of men's batterer treatment groups. *Research on Social Work Practice*, 11(6), 645-670.
[Random assignment-; control group-; sample size=104; problem: interesting study in that it evaluated group treatments for male batterers that provided men with "affective education [that] helps them to resolve their childhood traumas" as well as problem solving skills to end violent behavior, but outcomes measures were derived only from men's self-reports and therapist ratings.]
29. Yarbrough, D.N., & Blanton, P.W. (2000). Socio-demographic indicators of intervention program completion with the male court-referred perpetrator of partner abuse. *Journal of Criminal Justice*, 28(6), 517-526.
[Random assignment-; control group; sample size=286; problem: this purely observational study does not add much to what is already known about demographics of treatment completers vs. noncompleters]

Other Articles of Interest

30. Ehrensaft, M.K., Cohen, P., Brown, J., Smailes, E., Chen, H., Johnson, J.G. (2003). Intergenerational transmission of partner violence: A 20-year prospective study. *Journal of Consulting and Clinical Psychology*, 71(4), 741-753.
31. Gondolf, E.W. (2009). Implementing mental health treatment for batterer program participants: Interagency breakdowns and underlying issues. *Violence Against Women*, 15(6), 638-655.
32. Hamberger, L.D., Lohr, J.M., Gottlieb M. (2000). Predictors of treatment dropout from a spouse abuse abatement program. *Behavior Modification*, 24, 528-552.
33. Holtzworth-Munroe, A., Meehan, J.D., Herron, K., Rezman, U., Stuart, G.L. (2003). Do subtypes of martially violent men continue to differ over time? *Journal of Consulting and Clinical Psychology*, 71(4), 728-740.
34. Rosenberg, M.S. (2003). Voices from the group: Domestic violence offenders' experience of intervention. *Journal of Aggression, Maltreatment et Trauma*, 7(1-2), 305-317.
35. Smith, M.E. & Randall, E.J. (2007). Batterer intervention program: The victim's hope in ending the abuse and maintaining the relationship. *Issues in Mental Health Nursing*, 28, 1045-1063.
36. Taft, C.T. & Murphy, C.M. (2007). The working alliance in intervention for partner violence perpetrators: Recent research and theory. *Journal of Family Violence*, 22(1), 11-18.

APPENDIX C

Summaries of Higher-Quality Articles

APPENDIX C
Summaries of Higher-Quality Articles
[See Appendix E for copies of each of these articles]

1. Babcock, J.C., Green, C.E, Robie, C. (2004). Does batterers' treatment work? A meta-analytic review of domestic violence treatment. *Clinical Psychology Review*, 23(8), 1023-1053.

Design: Meta-analysis of 5 experimental and 17 quasi-experimental studies

Approaches studied: Duluth feminist psycho-educational model, cognitive behavior therapy (CBT) and other treatments such as couples therapy

Objective: To quantitatively summarize the findings to date (2004) on the effect of BIPs on violence recidivism

Methods: The investigators gathered published reports of BIP effectiveness studies in the academic literature using standard search methods. Studies were included if they: 1) had some form of comparison group of offenders and 2) relied on victim report or police record as the measure of recidivism (i.e. not offender self-report). The combined sample size was 1827 for the experimental studies and ~2030 for the quasi-experimental studies. The authors note that all of the quasi-experimental studies 'share the methodological problem of potentially "stacking the deck"' in favor of finding treatment effectiveness because they either compare findings for participants who completed treatment against those who dropped out (likely, a very different subgroup of offenders) or against a matched group of offenders who were not offered treatment or who were unwilling to attend treatment (also probably a different group of offenders).

Measures: Partner report of violence, police reports of rearrest

Results: Overall, there is a small positive effect of treatment on the chance of future violence, with treated offenders having a 40% chance of being successfully nonviolent compared to 35% for nontreated batterers, by partner report. This statistic means that 60% of treated batterers and 65% of nontreated batterers go on to reassault their victims. The authors found no statistically significant differences among the 3 treatment methods on later violence, meaning that none of the treatments studied was more effective at reducing violence than any other type of treatment.

Conclusion: The most widely available methods for treating violent batterers result in a 5% decrease in later violence toward victims. The authors note that this result is either a cause for celebration (in that, using U.S. prevalence statistics, this number equates to about 42,000 women per year no longer being battered—if all batterers attended a treatment program) or despair (given the costs associated with treatment provision and other "side effects" of unsuccessful treatment). The authors also note that, given the lack of strong findings for the effectiveness of treatment at all and, certainly, for the effectiveness of any particular treatment, states should not issue mandates limiting the range of treatment options for BIP programs.

Strengths: This is a methodologically sound study with a very large combined sample and good discussion of findings and of clinical and policy implications of the findings.

Limitations: None

2. Brannen, S.J., Rubin, A. (1996). Comparing the effectiveness of gender-specific and couples groups in a court-mandated spouse abuse treatment program. *Research on Social Work Practice*, 6, 405-424.

Design: Randomized comparative effectiveness trial

Approaches studied: Couples group intervention and gender-specific groups for batterers and victims

Setting: Research setting

Methods: Forty nine intact couples who indicated a desire to remain together were referred via a county court system. The majority of couples (67%) had been involved in relatively minor incidents of abuse and in 33%, the perpetrator had engaged in severe physical abuse such as punching, choking, kicking, use of a weapon. Couples were randomly assigned to a couples group or a gender-specific group intervention. The couples intervention used a CBT model designed to enable clients to accept personal responsibility for violent behavior and that included specific anger control techniques and focused on eliminating violence in the relationship. The men's gender-specific group included traditional Duluth Model components. The women's group was seen as supplemental to the perpetrators' group and focused on developing a sense of empowerment and strategies for safety.

In this study, an impressive and "elaborate safety net was established to ensure that none of the women were placed into a position of receiving further abuse as a result of their participation in the study" (article p. 412). These procedures are worthy of review by program developers and researchers who might be considering couples therapy as a treatment option.

Measures: Perpetrator and partner ratings of conflict resolution ability, level of violence, level of communication, marital satisfaction and recidivism, the latter measure confirmed by official police and court records.

Results: A significant decrease after intervention was found for the couples group on victim reports of low level abuse and severe physical abuse and this difference is mostly accounted for by the couples in the group in which the men had substance abuse problems. In other words, the couples intervention was particularly effective in reducing abuse in couples with husband substance abuse. Similarly, there was a decrease in abuse by substance abusing perpetrators in the gender-specific treatment as well that was not as large as the improvement for men in the couples treatment.

Recidivism at 6 months showed no difference between groups. There was no evidence to support the concern that victims in couples interventions experience more safety threats or incidents than victims in gender-specific groups.

Conclusion: Couples intervention may be especially effective for couples in which perpetrator substance abuse is an issue. Couples intervention does not appear to cause heightened safety risk for victims.

Strengths: Random assignment to treatment groups; strong safety context for victims involved in intervention.

Limitations: Lack of a control group.

3. Dunford, F.W. (2000). The San Diego Navy experiment: an assessment of intervention for men who assault their wives. *Journal of Consulting and Clinical Psychology*, 68(3), 468-476.

Design: Randomized controlled trial.

Approaches studied: Cognitive-behavioral men's group, cognitive-behavioral couples' group, rigorous monitoring group vs. control group in which men received no treatment and their wives received "stabilization and safety planning"

Objective: The purpose of the study was to experimentally evaluate the effectiveness of cognitive-behavioral interventions in different treatment settings for men who batter.

Setting: The Family Advocacy Center, a Navy agency responsible for the treatment of men who abuse their wives.

Methods: 861 married Navy couples in which active-duty husbands were substantiated as having physically assaulted their wives were randomly assigned to 4 groups: a cognitive-behavioral men's group, a cognitive-behavioral conjoint group (men and wives) with a communicational emphasis, a rigorous monitoring group that can be considered minimal treatment, and a control group with no treatment for the men and stabilization and safety measures for the women, that can be considered no-treatment.

Measures: Outcome measures included a self-reported measure assessing the number of incidents or episodes in which a victim or perpetrator reported being abused across 3 levels of abuse; abusive behaviors reported by respondents measured with the Modified Conflict Tactics Scale; official police and court records for all respondents; and the date of the first instance in which a repeat case of spouse assault occurred.

Victims and perpetrators were interviewed separately four times over the course of the experiment, at approximately 6-months intervals over the 18-month experimental period.

Results: The study found that no statistically significant differences on continuation of abuse between the 4 experimental groups using men's and women's reports of abuse and arrest records.

Conclusion: The cognitive-behavioral model, as implemented in this study via both men-only groups and couples groups, demonstrated little power to foster change in men receiving treatment for spouse abuse.

Strengths: Rigorous randomization, large sample size, high rate of completed interviews at extended follow-up

Limitations: Results probably cannot be generalized beyond the Navy population because of its special demographics.

4. Eckhardt, C.I., Murphy, C., Black, D., Suhr, L. (2006). Intervention programs for perpetrators of intimate partner violence; Conclusions from a clinical research perspective. *Public Health Reports*, 121, 389-381.

Design: Literature review

The authors conclude that while data regarding BIP effectiveness have improved over recent years, much is simply unknown about how such programs should be designed and how they should be applied in the field.

Approaches studied: Varies across studies reviewed.

Objective: To review the published empirical data on the effects of batterers intervention programs

Methods: Varies from study to study

Measures: The authors note that it is difficult to know what the most appropriate outcome measure is in batterers' intervention research. Looking at recidivism rates is problematic because rates of arrest are relatively infrequent for batterers (i.e., they probably engage in violent behavior far more often than they are arrested for it). Therefore, studies that do not have long follow up periods are unlikely to find differential outcomes for group vs. controls. They encourage researchers to use victim reports of psychological abuse (as well as acts of violence); however, these outcomes are more difficult gather.

Results: The authors report that evidence for the effectiveness of programs is very weak. Sample findings are:

- Between 40% to 60% of men mandated to BIP treatment either do not attend a group or drop out before finishing.
- Effects of treatment, where found, tend to be small.

- The more rigorous the research design, the smaller the effect size found (meaning that the studies that found less significant findings are probably more accurate).
- Some researchers have tried to compare various types of BIPs with each other, using rigorous research standards. Essentially, these studies found no difference or only small differences for the treatment groups compared to controls according to police reports of recidivism and partner reports.
- The few studies that directly compared traditional BIP treatments with couples therapy found no differences in outcomes between the groups. This either means that neither one is particularly effective or that couples therapy is as effective as traditional BIP treatments.
- Because most BIP studies are not well-designed or controlled, there is no way to rule out alternative explanations for studies that show a positive treatment effect.

Conclusion: The authors conclude that, given the above, “There are no interventions for partner violence perpetrators that approach [this] standard of being ‘empirically valid’, and it is debatable whether any intervention can [even] be labeled ‘empirically supported.’”

Recommendations: The authors suggest that it is time to develop BIP research methodologies similar to the methods used over the past few decades to study the differential effectiveness of psychotherapy modalities. Such studies would include:

- Sufficient number of participants to detect modest intervention effects (using statistical power analysis to determine needed sample sizes ahead of time)
- Careful screening of participants to make sure the participant group is relatively homogeneous
- Comparison of one or more well-described treatments with a manual that specifies in detail what the treatment involves
- Methods for measuring whether interventionists deliver the treatments as written
- At least one type of control group
- Random assignment to treatment and control arms of the study
- Multiple measures of outcome with, ideally, more than one reporter (i.e. police records, victim reports, clinician ratings, etc.)
- Data gatherers who are not involved in the delivery of treatments (to guard against bias)
- Detailed tracking strategies and incentives to reduce drop-outs and to insure that final data can be gathered even for the men who drop out.
- Sophisticated data analyses.

Strengths: This article is a good review of the literature and of factors that need to be considered in developing better research plans for the future.

Limitations: There is no clear description or listing of the studies reviewed, of inclusion criteria for studies included in the review, nor description of the search process.

5. Edleson, J.L., & Syers, M. (1990). Relative effectiveness of group treatments for men who batter. *Social Work Research and Abstracts*, 26(2), 10-18.

Design: Randomized comparative effectiveness trial.

Approaches studied: an education model delivered by trained “teachers” who provided information, a workbook and between-session assignments with little opportunity for discussion; a self-help model facilitated by a former batterer in which group members defined the topics covered but that always covered the topics of personal responsibility, a personal nonviolence plan, use of “time out” to diffuse tension, and the cycle of violence; and a combination of the two approaches. Each type of treatment was delivered in 2 intensities (12 sessions or 32 sessions).

Setting: Research setting

Methods: 283 men aged 17-57 who contacted the sponsoring agency were included in the study sample and were randomly assigned to one of the six treatment conditions described above. About one-third (N=102 or 38%) of the men were ordered to treatment by courts and the rest entered treatment voluntarily under social pressure. A total of 36 treatment groups were conducted over a 12 month period. All treatments were delivered in group a group format.

Measures: The main outcome variables considered were violence and threats of violence as reported by the men at beginning and end of treatment and by their partners at 6 months post-treatment (or by the men themselves if a partner could not be located).

Results: There were no significant differences found on any type of threats or violence between 12- and 32-week versions of treatment or between any of the treatment types. Participants in the self-help groups were more likely than participants in the other groups to have been violent at follow up, but these results were not significant.

Strengths: Random assignment to treatment groups; attempts to use partner reports of violence as a follow up outcome measure, attempt to use a rigorous design within a clinical agency.

Limitations: Lack of a control group. Since the study was confined to one setting only, findings are limited in generalizability. Significant attrition from groups occurred between intake and follow-up, thus reducing the possibility of finding significant results and the generalizability of results that were found.

6. Feder, L., Dugan, L., (2002). A test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward experiment. *Justice Quarterly*, 19(2), 343-375.

Design: Randomized controlled trial

Approaches studied: 26-week Duluth Model intervention

Objective: To attempt to answer the question, “Can courts effect change in spousal assault by mandating men who are convicted of misdemeanor domestic violence into a spouse abuse abatement program?”

Setting: Court system in Florida studying treatment provided in local BIP programs.

Methods: During a 5-month period, all men convicted of misdemeanor domestic violence in two courts in Broward Co., FL, were randomly assigned to an experimental group that received 26 weeks of group treatment from one of five local BIPs following the Duluth Model of treatment + one year probation or to a control group who received one year probation only. The final sample included 404 men.

Measures: Measurements used were offenders’ and victims’ surveys, attrition analysis of sample, and official records of rearrest. Offenders’ and victims’ surveys included an abbreviated version of the Inventory of Beliefs About Wife Beating Scale that assesses respondent’s view of the appropriateness of wife battering and the correctness of government intervening in such cases; a shortened Attitudes Toward Women Scale measuring men’s perceptions of the appropriate roles for women; criminalization of domestic violence; attitudes about partner’s responsibility; self-reported likelihood to hit partners again; and The Conflict Tactics Scale.

Results: About one-third of the ordered men failed to attend the intervention programs. There were no demonstrable positive effects of intervention on offenders’ attitudes, beliefs, or behaviors from participating in treatment groups. No differences were found between control and experimental groups in the likelihood of reoffending and being rearrested during the follow-up period. Twenty four percent of men in both the experimental and control groups were rearrested on one or more occasions during the year of probation.

Subanalyses provided the information that men who care little about the consequences of missing their court-mandated treatment sessions are also less concerned about the consequences of reoffending. This finding suggests that the men who attended all their treatment sessions would have avoided rearrest even without being mandated into the program. In other words, the men who completed treatment versus dropping out were a subgroup of men who were unlikely to reoffend anyway.

Conclusion: This study provides no evidence for the effectiveness of the Duluth Model of intervention.

Strengths: The study was conducted in a jurisdiction where men were closely monitored and sanctioned.

Limitations: Low response rate for victims, high turnover of research staff, insufficient sample to conduct analyses on the benefits of non-mandated counseling which was voluntarily attended by only 5 men in the total sample.

7. Feder, L., Wilson, D.B. (2005). A meta-analytic review of court-mandated batterer intervention programs: Can courts affect abusers' behavior? *Journal of Experimental Criminology*, 1, 239-262.

Design: Meta-analysis of 4 experimental and 6 quasi-experimental studies

Approaches studied: Psycho-educational feminist Duluth model all-male groups; cognitive behavioral all-male groups; one study that also assessed couples intervention groups and a rigorous monitoring-only intervention

Objective: To assess the effects of post-arrest mandated interventions (including pre-trial diversion programs) in reducing domestic violence offenders' future likelihood of re-assaulting through a synthesis of the available empirical literature

Methods: The investigators gathered published reports of BIP effectiveness studies in the academic literature using standard search methods. Studies were included if they: 1) used an experimental design (random assignment to groups + a control group) or a rigorous quasi-experimental design (ensured that the group being compared to the treatment group, although not randomly assigned, was equivalent on important factors to the treated group + used appropriate statistical methods); 2) interventions studied were court-mandated with the goal of reducing future re-assault behavior; 3) followed offenders for at least 6 months post treatment; and 4) used one or more objective measures of repeated violence (official or victim reports)

Measures: Victim reports and official police records

Results: Some support for modest benefits of BIPs is found when looking at official reports of arrests, but no effectiveness is found at all when looking at victim report measures. The authors note 4 strong concerns about the studies' findings. They believe the results of studies included in the meta-analysis are not generalizable to non-mandated batterers. Second, they believe there is a potential bias when official records are used as the outcome measure, due to victims' frequent unwillingness to file a complaint against the batterer or call the police. Third, the high rate of unavailability of victims across studies for treatment follow up assessment is problematic, they believe, and potentially biases studies toward finding positive results. Finally, the authors remind readers that using treatment drop outs as the comparison group is fraught with potential biases as well.

Conclusion: No clear effectiveness for any treatment method for court-mandated batterers was found. The authors recommend that the criminal justice system consider other types of interventions for addressing the problem of domestic violence and that such interventions be piloted and delivered via studies using an appropriate experimental design.

Strengths: This is a methodologically sound meta-analytic study.

Limitations: None

8. MacLeod, D., Pi, R., Smith, D., Rose-Goodwin, L. (2009). Batterer intervention systems in California. An evaluation. Judicial Council of California, Office of the Courts. Full text available online at: www.courtinfo.ca.gov/reference/documents/batterer-report.pdf.

Design: Program evaluation study that isolates specific components of the batterer intervention system to assess how differences in the system interventions affect outcomes for men who are in the system.

Approach studied: all state-certified 52-week BIPs and courts specialized procedures in California. BIPs reported employing educational models and skills training that included, at a minimum, elements of both the Duluth and cognitive-behavioral models. Programs tended to emphasize educational topics over skills training for batterers.

Objective: The purpose of the evaluation was to compare the efficacy of the justice system response across jurisdictions by looking at offender outcomes. Specifically, the study tried to determine: whether intervention impacts vary systematically across different jurisdictions; whether impacts vary systematically across BIPs within a jurisdiction; and whether program level variance accounts for differences in jurisdictional effects. Additionally, the study attempted to measure psychosocial changes in offenders resulting from program enrollment.

Methods: The study examined a sample of five jurisdictions in California and drew on a sample of approximately 1400 men enrolled in treatment programs across the five jurisdictions. The study took advantage of the fact that each jurisdiction managed its cases differently. Offender outcomes were measured by rates of program completion and rates of re-offense by offenders.

Measures: Attendance records for each offender enrolled in the study were analyzed to discern patterns in attendance, absences, and termination. The study identified offender characteristics that were strongly correlated with program termination and completion. Those risk factors were used as control variables in analyses that were used to answer the main questions of the study.

Results: The evaluation found that the strongest predictor of rearrest following intake in a BIP was the individual characteristics of the offenders rather than the characteristics of jurisdiction or of the BIPs in which offenders were enrolled. Men who were more educated, older, had shorter criminal histories and did not display signs of drug or alcohol dependence had a lower likelihood of rearrest independent of the kind of treatment they received.

Conclusion: Individual characteristics are more salient in predicting program completion and re-offense than the type of treatment. Thee authors recommend enhanced risk and needs assessments at intake to improve offender treatment and outcomes and the greater availability of drug and alcohol treatment concurrent with BIP treatment for offenders.

Strengths: This is a unique large scale cross-jurisdiction evaluation of BIP outcomes that led to statistically robust findings. Difficulties and limitations of measurement are carefully delineated in the report. Both research and policy implications are carefully discussed.

Limitations: None

9. Morrel, T.M., Elliot, J.D., Murphy, C.M., Taft, C.T. (2003). Cognitive Behavioral and Supportive Group treatments for partner-violent men. *Behavior Therapy*, 34, 77-95.

Design: Comparative effectiveness study with quasi-random assignment to treatment (see the article (p.81 of article: Assignment to Conditions for details of quasi-randomization procedure).

Approaches studied: Cognitive-behavior therapy (CBT) and supportive group therapy for men.

Objective: To determine whether a structured, skills training group based on the principles of CBT was more effective than unstructured, supportive group therapy in reducing rates of physical and psychological abuse and in affecting secondary treatment targets that may confer risk for continued problems with abuse.

Setting: A community domestic violence agency in Maryland

Methods: Eighty six men seeking group treatment for partner-abusive behavior were systematically assigned to cognitive-behavioral group therapy (CBT) and to a relatively unstructured supportive group therapy (ST) at a community center.

Measures: Criminal recidivism, aggression reported by partners, global impression of change, communication behaviors, readiness to change, self-esteem and self-efficacy. Measurements were based on partner reports at 6 months and official reports of criminal recidivism at 2 to 3 years.

Results: There were no significant treatment differences between CBT and ST based on data from both partner reports of criminal recidivism and criminal data. Both CBT and ST were associated with significant reductions in physical assault, psychological aggression, injuries and sexual coercion and with increases in self-esteem and self-efficacy,

Conclusion: The study failed to demonstrate an added benefit of a CBT group intervention over the effect of a minimal supportive group treatment experience for men who volunteered for batterer treatment.

Strengths: Design and analysis strengths include careful consideration to treatment dropout, examination of treatment adherence and control for therapist effects. Outcome data were collected from multiple sources and for a long period of time after end of treatment

Limitations: Although treatment assignment was systematic, it was not random and this may limit the validity of the findings.

10. O’Leary, D.K., Heyman, R.E., Neiding, P.H. (1999). Treatment of wife abuse: a comparison of gender-specific and conjoint approaches. *Behavior Therapy*, 30, 475-505.

Design: Quasi-randomized comparative effectiveness study

Approaches studied: Two therapy formats for couples with repeated acts of husband-to-wife physical aggression: a gender-specific treatment (men-only and women-only) group therapy and conjoint couples therapy, both therapy types based on a cognitive-behavioral model; in the gender-specific groups, men were held responsible for aggression; in the conjoint groups, both men and women were considered as sharing responsibility for reducing marital discord.

Objective: To provide a comparison of the effectiveness of 2 treatment approaches focusing on the reduction of psychological and physical aggression, in a self-referring, martially intact, physically aggressive sample. The study also aimed to test concerns about the safety of and other controversies regarding couples therapy when domestic violence is present.

Setting: Unclear, but appears to be a research laboratory setting

Methods: 75 intact volunteer couples were assigned to either a gender-specific treatment condition (male and female groups meeting separately) or a conjoint 14-week group therapy for psychological and physical aggression. To participate, couples reported 2 or more acts of husband-to-wife aggression in the past year that did not result in injuries needing medical attention. Couples had to be willing to be randomly assigned to either treatment modality; wives, when interviewed separately, had to report they would be comfortable being in conjoint treatment with their husbands, among other inclusion criteria. Quasi-randomization procedure: eligible couples were placed on a waiting list and, when 6 to 8 couples qualified, a new group was started, alternating between gender-specific treatment and conjoint treatment. Forty couples were assigned to conjoint therapy and 30 to gender-specific therapy. Both modalities lasted 14 weeks.

Measures: Self-report measures were administered at pretreatment, posttreatment, and 1 year follow-up, and included: frequency of functional and verbally and physically abusive tactics used during marital conflict; dominance/isolation, fear of spouse; attribution of responsibility; depression; dyadic adjustment; fear and/or aggression due to treatment sessions; and participant satisfaction.

Results: Across treatment type, men reduced severe physical aggression by 51%, moderate physical aggression by 55%, and psychological aggression by 47%. Only one-fourth of men were completely violence-free at 1-year follow-up, but two-thirds of men

maintained cessation of severe aggression. Significant improvements at post-treatment and follow-up were found for both spouses' marital adjustment, wives' depression, and husbands' taking responsibility for aggression, again independent of treatment type.

Regarding women's safety, couple's arguments regarding issues discussed in treatment led to physical aggression in only 2% of sessions for both groups, with no difference between the groups on this measure. There was no evidence that women were more afraid to express themselves in couples therapy than in gender-specific groups. Both male and female participants were highly satisfied with both forms of treatment, with no differences between the treatment groups.

Conclusion: Both gender-specific and conjoint treatment of volunteer couples resulted in significant decreases in aggression and other personal and marital improvements over time. Neither treatment was superior to the other in terms of safety and effectiveness. The concern that women's risk of victimization would increase in conjoint therapy was not supported.

Strengths: This is one of the few studies to explore the comparative effectiveness of conjoint and gender-specific group therapy. It is very well-designed and investigators tested for therapist adherence to treatment protocols and other potentially confounding factors.

Limitations: The results found with this volunteer sample cannot be generalized to a sample with perpetrators receiving mandated referral to treatment from the court system or to couples in which the woman would be afraid to be in conjoint counseling with her husband.

11. Saunders, D.G. (1996). Feminist-Cognitive-Behavioral and Process-Psychodynamic treatments for men who batter: Interactions of abuser traits and treatment models. <i>Violence and Victims</i>, 11(4), 393-414.

Design: Randomized comparative effectiveness study.

Approaches studied: Feminist-cognitive-behavioral treatment and process-psychodynamic group treatments.

Objective: To improve on previous BIP evaluations by obtaining a higher rate of response during follow-up and by ensuring that the treatments studied were applied according to their stated goals. The investigator hypothesized that each of the two treatments would be differentially effective for batterers with specific traits.

Setting: Community-based domestic violence program

Methods: 213 men were randomly assigned to one of the treatment interventions. Most participants (76%) were referred by a deferred prosecution program or probation department following prosecution, while most of the others volunteered for treatment as a

result of “social pressure.” The two treatments compared were both offered in close-ended groups of 20 weekly sessions lasting 2.5 hours each.

Measures: Recidivism measured by the victim reports at 3 to 54 months after treatment and supplemented by men’s reports and arrest records; psychological abuse, level of fear for victims, general changes in men and use of conflict resolution methods.

Results: No significant differences were found between the two programs on victim reports of violence, fear, general changes in their partners, or relationship equality at 22 or more months after treatment. As hypothesized, results showed that offenders with dependent personalities had significantly lower rates of recidivism in the process-psychodynamic groups, while those with antisocial personalities had lower recidivism rates in the structured, feminist-cognitive-behavioral groups. Batterers with substance abuse potential and hypomania also had lower recidivism in the feminist-cognitive-behavioral treatment condition.

Conclusion: Personality styles and disorders of batterers interacted with the type of treatment received. There may not be a “one size fits all” approach to batterer treatment.

Strengths: This study was rigorously designed and implemented. It assessed the effects of treatment integrity as well as the effect of drop out rate on potential bias in study results. The study relied on stringent measures of recidivism. Additionally, the study demonstrated successful long-term follow-up.

Limitations: None

12. Stover, C.S., Meadows, A.M., Kaufman, J. (2009). Interventions for intimate partner violence: review and implications for evidence-based practice. *Professional Psychology: Research and Practice*, 40(3), 223-233.

Design: Literature review

Approaches studied: Mandatory arrest, Duluth model group treatment, group cognitive behavioral therapy (CBT) or combined CBT-psychoeducation intervention for batterers, and couples intervention. The review also includes studies of victim and child witness interventions, the results of which are not reported on here.

Objective: To survey available intimate partner violence treatment studies with randomized case assignment and at least 20 participants per group.

Methods: A literature search was conducted using MEDLINE and PsycINFO databases using accepted methods. To be included, studies had to meet the following criteria: 1) used a randomized controlled research design; 2) had at least 20 participants per treatment group; and 3) included recidivism or measures of violence severity as outcomes (except for couples intervention studies which were, in general, poorly designed and which could compare one treatment against another without a control group).

Measures: Police and victim reports of violence

Results: One-third of batterers treated in any of the modalities tested will have a new episode of violence within 6 months of end of treatment, with no difference among treatment modalities. Recidivism rates were notably higher when measured by victim reports compared to police reports, but there was a high rate of missing victim data in most studies, calling into question the overall results. The one well-designed couples intervention study from 1988 that included a multi-couple group intervention compared to individual couples intervention found a 20% recidivism rate at 6 month follow up for both (lower than most men-only treatment results), but attrition from the groups was so high that results are in question.

Conclusion: Rigorous evaluations of group treatments for batterers show minimal or no impact compared to mandatory arrest alone. There are preliminary data to support the potential effectiveness of couples interventions, especially for those where the batterer has an alcohol and/or substance abuse disorder. The authors conclude that there is "...a lack of research evidence for the broad, long-term effectiveness of many of the most common treatments (...) including the Duluth model for perpetrators" (p. 231) and note that "policies requiring specific treatment approaches for all male batterers are not effective" (p. 231).

Strengths: This is a literature review based on a strong search methodology that describes only the most rigorous published studies. The authors make specific policy and treatment development recommendations.

Limitations: None

13. Stuart, G.L, Temple, J.R, Moore, T.M. (2007). Improving batterer intervention programs through theory-based research. *JAMA*, 298(5), 560-562.

Design: Literature review

Approaches studied: N/A

Objective: To inform program administrators, policy makers and researchers by describing briefly what is known about the efficacy of BIPs, describing reasons for the ineffectiveness of current BIPs and making recommendations for improving effectiveness of programs

Methods: N/A

Measures: Not discussed

Results: "Numerous studies, qualitative reviews, and meta-analyses have repeatedly arrived at a similar conclusion: batter intervention programs have a small, often nonsignificant effect in reducing partner violence" (article p. 560). Reasons for this ineffectiveness are hypothesized to be: 1) batterers are usually court-mandated and may

be unwilling or unmotivated to accept responsibility for being violent; 2) BIPs receive inadequate funding and, therefore, have limited resources and often employ overworked clinicians who lack professional counseling degrees; 3) interventions are seldom tailored to clients' needs; and 4) programs were rushed into use and mandated by states before their effectiveness was rigorously evaluated.

Conclusion: Recommendations are to: 1) make use of motivational theories and strategies in programs, such as the Transtheoretical Model of Behavior Change and Motivational Interviewing; 2) tailor treatment to meet the needs of batterer subgroups; 3) Include substance abuse treatment as part of BIP services; and 4) consider and evaluate the use of couples treatment for carefully-selected batterer-victim dyads.

Strengths: This is a brief but very strong literature review of the major findings in BIP research with well thought-out recommendations based on the gaps in the literature.

Limitations: None

14. Taylor, B.G., Davis, R.C., Maxwell, C.D. (2001). The effects of a group batterer treatment program: A randomized experiment in Brooklyn. *Justice Quarterly*, 18(1), 171-201.

Design: Randomized controlled trial

Approach studied: 40-hour Duluth model program

Objective: To test batterer treatment using an experimental design that randomly assigns court-mandated batterers to treatment or control conditions and to address methodological problems from prior research, including disentanglement of the effects of treatment from sample selection effects.

Setting: A batterer treatment program in New York City, in conjunction with the county court.

Methods: 376 male criminal court defendants charged with assaulting their female partners were randomly assigned to a 40-hour BIP Duluth model group or a control group that had to complete 40 hours of community service that included cleaning local parks and public buildings. In order for a defendant to be included in the study, all parties--including the defendants, the judge and prosecutor--had to agree to that the defendant would/could participate in batterer treatment if he was assigned to that condition.

Measures: Recidivism reports were collected from multiple sources including arrest reports, crime complaints, and victims' reports of violence. Four recidivism measures were constructed: prevalence, rate or frequency of failures, severity, and time to the first failure. Follow-up measurements were collected at 6- and 12-month post-sentencing.

Results: Men in the treatment group showed significantly lower recidivism from official records at 6-month follow up (treatment group=16% recidivism; control=38%) and 12-month follow up (treatment=28%; control=55%). Victims' reports indicated much higher recidivism for all men (6-mo: treatment group=67%; control=90%; 12-mo: treatment group=46%; control=99%). Although these latter numbers appear to indicate big differences between the groups, the number of victims reached at those time points was so small that the results are not statistically significant and cannot be considered to represent true differences between the groups.

Investigators also looked at "time to first official failure", a measure of the time interval before batterers were rearrested for assault on their intimate partner. Time to first failure was significantly longer for men in the intervention group compared to the control group, thus creating "a consistent period of greater safety for victims during the first year of follow up" (p. 193).

Conclusion: Results of this study show some support for the effectiveness of a Duluth model intervention compared to a control group in, at least, extending the length of time to reassault for court-mandated batterers.

Strengths: Strong experimental design.

Limitations: The final sample may be an unrepresentative sample of court-mandated batterers, as only 373 of more than 11,000 sentenced batterers were included in the study based on inclusion criteria. Low response rate in victims is another limitation. In addition, there were 53 cases assigned to the control group that were reassigned by judges to the treatment group after the fact, thus potentially significantly influencing the study findings.

APPENDIX D

Summaries of Medium-Quality Articles

APPENDIX D

Summaries of Mid-Quality Articles

15. Dutton, D.G., Bodnarchuk, M., Kropp, R., Hart, S.D., Ogloff, J.R.P. (1997). Wife assault treatment and criminal recidivism: An 11-year follow-up. <i>International Journal of Offender Therapy and Comparative Criminology</i>, 41, 9-23.
--

Design: Observational

Approach studied: unspecified “anger management” or “spousal assault treatment”

Objective: To assess over a long follow up period the results of treatment in terms of the Prochaska et al. (1992) model of stages of change.

Setting: A batterers intervention program in Vancouver, British Columbia.

Methods: The sample included 446 voluntary and court-referred offenders assessed at this program over a 10-year period from 1982-1992 whose criminal records were available from a national database in 1993. The sample was divided into the following categories and the outcomes for each were described: Completers (attended at least 12 of 16 sessions), Noncompleters (attended fewer than 12 sessions; average was 5.3 sessions); and No Shows (referred to the program but did not attend the intake interview). Another group called Rejects (completed intake but either were not willing to participate or did not meet other inclusion criteria) was identified, but was not included in the analysis.

Measures: Reassault determined by court records.

Results: Completers were more educated, more likely to be employed, more likely to be currently in a relationship at the time of intake and had lower precontact rates of criminal offenses, violent crimes and assaults than Noncompleters. There were no differences in these variables between Completers and No Shows; thus, men who never showed up for treatment were no different demographically or in their criminal record than those men who went on to complete the mandated BIP intervention.

During the follow up period, there was no difference in reassault between Completers and Noncompleters. Most men exhibited either zero or one repeated assault (that was recorded in the crime database), with a very small number of men spread across the categorized groups that committed a large number of reassaults.

Conclusion: The single best predictor of future partner assault is level of past partner assault, independent of treatment completion or noncompletion.

Strengths: This study followed a large sample of batterers over a significant period of time.

Limitations: Type of treatment received is not clearly specified. The study did not have

enough participants to allow for a valid test of the main objective (does treatment shorten the “assaultive career” of batterers).

16. Eckhardt, C., Holtzworth-Munroe, A., Norlander, B., Sibley, A, Cahill, M. (2008). Readiness to change, partner violence subtypes, and treatment outcomes among men in treatment for partner assault. *Violence and Victims*, 23(4), 446-475.

Design: Observational

Approaches studied: conventional BIP programs in the community, although the purpose of the study was not to determine the effectiveness of these programs per se.

Objective: To determine whether pre-BIP readiness to change and the presence of partner violence subtypes predicted completion of the BIP program, criminal recidivism, and post-adjudication partner violence at 6 months post intervention.

Setting: Family Violence Court in Dallas

Methods: The sample consisted of 199 court-mandated convicted male offenders who met eligibility criteria and kept their initial BIP program appointment.

Measures: Analyses were based on data from a pre-BIP interview with men and their partners and reviews of criminal justice outcomes at 6 months post-intervention follow-up. Outcome measures were: BIP completion; rearrests from official records; self- or partner-reported partner violence recidivism.

Other variables of interest were: stage of readiness to change; partner violence typology subtype, classified as family only, antisocial, borderline, or dependent; alcohol use; consequences related to drug-associated problems; automatic thoughts associated with hostility and anger arousal; endorsement of the appropriateness of use of violence in close relationships; attitudes toward women; frequency and type of anger responses.

Results: 40% of the sample did not complete BIP. Four readiness-to-change groups were identified; 76% of men had change-resistant profiles and the majority had little to no motivation to change behavior. Participants in the study belonged to four typological subtypes: family only, low-level antisocial, borderline/dysphoric, and generally violent/antisocial. BIP completion was predicted by violence subtype with the borderline/dysphoric and generally violent/antisocial types more likely to drop out. BIP completion was not predicted by readiness to change profiles. Rearrested men were more likely to belong to the borderline/dysphoric and generally violent/antisocial types.

Conclusion: Offenders in the study were not uniform on many important dimensions that may predict BIP completion and rearrest. The partner violence subtype construct may be useful in planning treatment.

Strengths: The study raises significant issues about the relevance of a general BIP

approaches, given the demonstrated selective influence of personality subtype on program engagement.

Limitations: Much of the data are self-reported by the offenders; only a small subset of women partners was available to provide corroborating information for their partners' self-report.

17. Gondolf, E.W. (1999). A comparison of four batterer intervention systems: Do court referral, program length, and service matter? *Journal of Interpersonal Violence*, 14(1), 41-61.

Design: Observational multisite evaluation

Approaches studied: Traditional services based on cognitive-behavioral/feminist psychoeducational approaches and traditional services plus additional services (e.g., in-house alcohol treatment or referral for alcohol treatment.)

Objective: To address some of the conceptual and methodological limitations of other studies and to further the research on the relative effectiveness of different batterer intervention systems

Setting: Well-established BIP programs in 4 U.S. cities

Methods: Four geographically distinct batterer intervention systems were selected for comparison of their differences along three components: court referral; program duration (3, 6, and 9 months); and presence or absence of additional services. At each site, the first 20-25 men appearing for program intake at the beginning of each month and who accepted to participate in the research became part of the sample until a total of 210 participants were recruited at each site. Final sample size was 840 men.

Measures: The primary outcome was reassault rates reported by women partners during a 15-month follow-up. Additional outcomes were controlling behaviors, verbal abuse, and threats and women's overall sense of safety and well-being.

Results: Rates of reassault and rates of other outcomes were relatively similar across sites at follow-up despite differences in batterer demographics, program format, and jurisdiction. Severe reassault was significantly lower for the longest and most comprehensive program.

Conclusion: There were no differences in outcomes across the range of programs investigated. The authors conclude that "differing intervention systems that conform to fundamental standards can achieve similar outcomes."

Strengths: This study attempted a useful comparison of program outcomes across geographical sites and across programs sharing fundamental essentials yet offering a range of different services.

Limitations: The selection of sites may have introduced significant confounders. Program content was not rigorously determined or measured. Due to its design, the study cannot hypothesize which factors explain the findings or whether results found are due to program effects or other factors or are comparable to or different from reassault rates for non-program attendees.

18. Gondolf, E.W. (2000). A 30-month follow-up of court-referred batterers in four cities. *International Journal of Offender Therapy and Comparative Criminology*, 44(1), 11-128.

Design: 30-month follow-up to an observational multi-site evaluation

Approach studied: Traditional services based on cognitive-behavioral/feminist psychoeducational approaches and traditional services plus additional services (see Gondolf, 1999, above for fuller description)

Objective: To complete long-term follow up (2 years after program intake) of court-referred batterers who were referred to a BIP program.

Setting: 4 well-established BIP programs

Methods: Follow-up data were collected by telephone interviews with males and their partners at 22 to 23 months after intake and at 30 months after program intake.

Measures: The primary variable of interest was reassault, measured by women's reports, of conflicts, physical aggression, the nature of battering injuries and medical assistance received for those. Other variables included: other abuse reported by women and women's subjective appraisal of overall well-being and safety.

Results: The outcomes across sites were the same as for the 15 month follow-up reported in Gondolf, 1999, above. There were no significant differences on rearrest rates or on the other outcome variables across the four locations. Cumulative reassault rates for all men who entered the program (including those that dropped out) varied from 34% to 47%. According to partner reports, 41% of the men reassaulted their partners during the 30-month follow-up. Analysis of reassault trends showed that there was only a 7% to 8% increase in reassault rates between 15 and 30 months after program intake. About 83% of first-time reassaults occurred during the first 15 months.

In respect to repeated reassault, 21% of men repeatedly reassaulted their partners over the 30-month period and those 21% were responsible for 60% of injuries counted. Between 15 and 30 months from intake about 80% of men had not reassaulted their partners. Other forms of abuse followed the trends of reassault. The majority of women felt better off and felt safe at the 30-month follow-up (an increase from 3% to 10%). There were no differences in re-assault rates across sites.

Conclusions: Most of initial reassaults after a BIP program intake occurred within the first 6 months and then progressively decrease in time. The author concludes that the trends observed are encouraging and life of the majority of partners seems to improve based on their subjective ratings.

Strengths: Follow-up response rates were high and this factor increases the value of the analyses. Drop-out effects were accounted for and comparative analyses of drop-out vs completers were performed.

Limitations: The non-experimental nature of the design does not allow for extrapolation of data to other populations and circumstances and does not allow for advancement of any hypothesis about the underlying mechanisms responsible for the results observed.

19. Gondolf, E.W. (2004). Evaluating batterer counseling programs: A difficult task showing some effects and implications. *Aggression and Violent Behavior*, 9, 605-631.

Design: Observational comparative evaluation design

Approach studied: Gender-based cognitive behavioral treatment with substantive site differences in structure and context

Objective: The purpose of the study was to address some of the conceptual and methodological shortcomings of previous BIP effectiveness research.

Setting: Four “well-established” BIP programs in four major U.S. cities

Methods: The evaluation involved a 4-year follow up, starting at program intake, with 840 court-referred male batterers and their female partners.

Measures: The main outcome was reassault based on victim report and backed up by analysis of police reports and men’s self-report.

Results: A 49% reassault rate was shown at 4 years across programs. The majority of reassaults occurred within 6 months from intake and the incidence of new assaults decreased over time. At the 4-year follow up, fewer than 10% of the men had assaulted their previous or current partners within the past year; over two thirds of the women said their quality of life had improved at 4 years and 85% reported feeling very safe.

Conclusion: The investigator concludes that there is evidence that a gender-based cognitive behavioral program “seems to be appropriate for the majority of men” and that such programs help batterers “stop their assaultive behavior and reduce their abuse in general.” These conclusions are reached, however, without comparing this approach to any other and without specifying attrition rates clearly.

Strengths: This evaluation has a large sample size, multiple sites, and sophisticated

measurement and statistical procedures.

Limitations: Lack of random assignment to treatment condition. The author improperly describes a “program effect” by comparing outcomes for those who completed the program compared to those who enrolled but dropped out, although he does make a case for using a statistical procedure called “propensity score analysis” for doing so.

20. Gordon, J.A., Moriarty, L.J. (2003). The effects of domestic violence batterer treatment on domestic violence recidivism. The Chesterfield County Experience. *Criminal Justice and Behavior*, 30(1), 118-134.

Design: Quasi-experimental with non-equivalent control group.

Approach studied: 20- and 24-week group Duluth Model feminist psychoeducational programs

Objective: The purpose of the study was to determine the influence of treatment on the recidivism rate of domestic violence offenders and to determine demographic characteristics associated with recidivism.

Setting: A county court system and 2 contracted BIP agencies in the community

Methods: The sample consisted of 248 male domestic violence offenders sentenced to Community Corrections Services in Chesterfield County, VA, between January and December 1999; 132 of the men were court-ordered to attend domestic violence treatment, while 116 men who received no mandatory treatment comprised the (non-randomized) control group.

Measures: The numbers of rearrests and reconvictions were the main outcome variables, collected from the VA Criminal Information Network after a follow-up period of at least one year.

Results: There were no differences in likelihood of rearrest or reconviction for offenders court-ordered into treatment compared to those who had not been ordered to treatment. Within the group that received treatment, the number of sessions received and the successful completion of the program were associated with reduced likelihood of rearrest and reconviction.

Conclusion: Offenders who received mandatory treatment did not show a decrease in recidivism after one year compared to offenders who did not receive treatment.

Strengths: Acknowledges the non-equivalency of the two groups compared and discusses the potential effect of this issue.

Limitations: Non-equivalent comparison groups. There likely was a reason for some men to be ordered to treatment and other men not to be, so the two groups being compared

were likely to have had preexisting differences not related to treatment. Results reported, including the observed association between sessions/program completion and reduced recidivism cannot be considered to be an effect of the BIP treatment.

21. Saunders, D.G. (2008). Group interventions for men who batter: A summary of program descriptions and research. *Violence and Victims*, 23(2), 156-172.

Design: Literature review

Approaches studied: Varies depending on the study reviewed; all programs reviewed had a treatment component of some kind; studies involving purely criminal justice interventions were not included.

Objective: To summarize recent research (through 2008) on all-male group interventions for men who batter, including the major components of programs, what is known about treatment effectiveness, and methods for enhancing treatment motivation and culturally competent practice.

Methods: A literature search was conducted that resulted in more than 35 program effectiveness studies that are reviewed.

Measures: Vary from report to report.

Results: Reducing attrition by increasing motivation of batterers participating in programs is of major importance. Several methods for doing so are described, including marathon orientation groups, culturally-tailored interventions, and motivational enhancement (a brief form of motivational interviewing). Authors briefly describe the few culturally competent interventions available in the literature. Approximately one third of victims report reassault within one year, by victim reports, across all types of programs. A promising avenue for future research is matching of offender type with type of treatment.

Conclusion: There is little well-designed empirical evidence to support the effectiveness of BIPs.

Strengths: This is an exhaustive review that highlights the major issues and challenges of BIP effectiveness research.

Limitations: A table describing the common characteristics of studies reviewed would make the article more clear.

22. Snow Jones, A., D'Agostino, R.B., Gondolf, E.W., Heckert, A. (2004). Assessing the effect of batterer program completion on reassault using propensity scores. *Journal of Interpersonal Violence*, 19(9), 1002-1020.

Design: Additional analysis of data from a previous multi-site study (see Gondolf 1999)

Approaches studied: N/A.

Objective: To address the concern of high attrition rates in BIP programs and to begin to answer the questions, “If we can reduce BIP program dropout, will there be a reduction in reassault?” and “Is there a significant effect of a greater or full dose of treatment?” [NOTE: This is not an actual treatment effectiveness study (although the authors make unfounded conclusions about treatment effectiveness based on their data), but a study of variables that predict treatment completion or drop out and the associations of completion status with later reassault.]

Setting: Data from 3 of the 4 sites described in Gondolf 1999 and 2000 were analyzed.

Methods: Using propensity score analysis, investigators estimated the probability of completing a BIP program, based on observable characteristics of participants. Propensity scores are computed using a statistical procedure that matches participants in a study using observed characteristics and then predicts the outcome of a target variable from the score. Using propensity scores to analyze data from the previous study, investigators derived a method of predicting subtypes of offenders and their likelihood of completing BIP treatment.

Measures: Personality, psychopathology and alcohol use; program completion; reassault.

Results: At all but one propensity level, completers were less likely to reassault when compared to program drop outs (26% vs. 39%, respectively) and this finding holds true for completion of any program, regardless of length. Men who enrolled voluntarily in treatment showed higher reassault rates for both drop outs and completers than men who were court-mandated (volunteers: 51% drop outs, 48% completers; mandated: 38% drop outs, 21% completers).

Conclusion: The authors conclude, inappropriately, that their findings are stronger than those derived from experimental studies with regard to the effect of treatment on reassault.

Strengths: The study uses a sophisticated statistical analysis to predict who may and may not drop out of BIP treatment. Results may lend themselves to improvements in retention strategies for programs.

Limitations: The lack of a control group in this sample limits the ability to link the treatment itself with later reassault rates. Other studies have found that BIP drop outs tend to show characteristics associated with reassault in the general batterer population, so that higher reassault rates for the drop outs and lower rates for completers in this study may not be related to treatment characteristics, but to individual characteristics of the participants.

23. Snow Jones, A., Gondolf, E.W. (2001). Time-varying risk factors for reassault among batterer program participants. *Journal of Family Violence*, 16(1), 345-359.

Design: Additional analysis of data from a previous multi-site study (see Gondolf 1999)

Approaches studied: N/A

Objective: To extend previous batterer research by using a dynamic model of reassault that includes both time-varying (situational, psychological) characteristics that may be risk factors for reassault as well as time-invariant (personality, sociodemographic and prior behavior) characteristics

Setting: Data from the 4 sites described in Gondolf 1999 and 2000 were analyzed.

Methods: Data collected at five points at 3-month intervals from a subset of 308 men in BIP treatment who were court-mandated (82% of the whole sample) and their partners were examined for time-varying situational and behavioral risk factors and time-invariant individual characteristics in their association with reassault events.

Measures: The outcome variable at 1-year follow-up was reassault rates based on partners' report. *Time-invariant* variables were ethnicity, age at intake, education, personality and behavior at intake, exposure to bad parental behavior. *Time-varying* variables were: unemployment during follow-up interval, drinking behavior, frequency of drunkenness, and help-seeking behavior.

Results: The time-varying behavioral characteristic of alcohol abuse (any drunkenness and high frequency of drunkenness after intake) was associated with the highest risk of reassault. At least one drunken episode during the follow up period was associated with a 3.5 times higher risk for reassault compared to the non-drunken group. Those who drank almost daily were 16 times more likely to assault than those who were not.

Two time-invariant individual characteristics were also positively and statistically significant associated with reassault: history of non-DV arrest and evidence of severe psychopathology at intake. No specific personality traits or types measured appear to be risk markers.

Conclusion: Findings suggest that batterers' drinking behavior may be a strongly predictive indicator of risk for reassault. Assessment of potential danger at intake may need to include measures of alcohol use and time-varying measures should be assessed along with the more usual assessment of time-invariant measures.

Strengths: Good correlational analyses with clinical sense and practical implications.

Limitations: No possibility of speculating on underlying mechanisms and no generalizability to the whole batterer population.

24. Taft, C.T., Murphy, C.M., King, D.W., Musser, P.H., DeDeyn, J.M. (2003). Process and treatment adherence factors in group cognitive-behavioral therapy for partner violent men. *Journal of Consulting and Clinical Psychology*, 71(4), 812-820.

Design: Observational study

Approach studied: Cognitive-behavioral group treatment in a 16-week closed-group format with a motivational enhancement component.

Objective: To elucidate process and adherence factors that may promote active change during the course of a 16-week cognitive behavioral group treatment program for partner violent men. Major hypotheses were that working alliance, group cohesion, session attendance and homework compliance would predict physical and psychological abuse at 6-month follow up.

Setting: A domestic violence treatment center in Maryland.

Methods: A sample of 107 men in treatment for intimate partner abuse perpetration was assessed over a one year period; 88% of the sample was court-mandated. The study examined process and treatment adherence factors as predictors of partner abuse following participation in the CBT group program. The treatment was divided into a sequence of four components aimed at enhancing motivation to change and providing self-regulation skills for and relationship alternatives to abusive behavior.

Measures: Outcome measures were: the strength of the working alliance between clients and therapists; group cohesion; homework compliance; session attendance; and abusive behavior, the latter assessed from partner reports. Statistical analyses used multilevel modeling to determine the relationship between the predictors and outcomes represented by collateral partner reports of abuse.

Results: Therapist working alliance ratings were the strongest predictor of outcome measured as physical and psychological abuse at 6-months follow-up. Client perceptions of the strength of the therapist-client alliance ratings were not related to outcomes. Group cohesion ratings and homework compliance predicted psychological abuse.

Conclusion: A supportive and collaborative therapeutic environment and a high level of group cohesion during treatment may be beneficial in helping partner violent men change abusive and violent behaviors.

Strengths: The research design was driven by predetermined hypotheses; measures and analyses were appropriate to the questions being studied.

Limitations: The study sample was limited to one program in one location. Results may not be generalizable to other locations or populations.

25. Whitaker, D. J. & Nolon, P.H. (2009). Advancing Interventions for Perpetrators of Physical Partner Violence: Batterer Intervention Programs and Beyond. In D. J. Whitaker and J. R. Lutzker, *Preventing partner violence: Research and evidence-based intervention strategies*. Washington, DC: American Psychological Association, pp. 169-192.

Design: Literature review

Approaches studied: Good description of available treatment approaches

Objective: To review intervention group and individual approaches for intimate partner violence that focus on the perpetrators of IPV.

Methods: N/A

Measures: Not discussed

Results: “The strongest evidence for BIPs’ effectiveness comes from the least rigorous studies” (article p. 171). Also: “There is little empirical evidence to support” the mandating by states of particular BIP approaches that emphasize patriarchy as a cause of violence and require group feminist-psychoeducational and/or cognitive behavioral treatment as the only acceptable and state-certifiable mode of BIP treatment. Discusses the current lack of and need for tailoring of interventions for batterer subgroups, including cultural subgroups and subgroups with alcohol and substance abuse; and for addressing and intervening with female perpetrators of partner violence.

Strengths: This review that addresses some issues that other reviews do not (i.e. women perpetrators of IPV).

Limitations: A table showing details of studies, common components and outcomes would make this article easier to synthesize.

APPENDIX E

Copies of Higher-Quality Articles



Pergamon

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

CLINICAL
PSYCHOLOGY
REVIEW

Clinical Psychology Review 23 (2004) 1023–1053

Does batterers' treatment work? A meta-analytic review of domestic violence treatment

Julia C. Babcock^{a,*}, Charles E. Green^a, Chet Robie^b

^a*Department of Psychology, University of Houston, 126 Heyne Building, Houston, TX 77204-5022, USA*

^b*College of Business Administration, Niagara University, Niagara Falls, NY, USA*

Accepted 29 July 2002

Abstract

This meta-analytic review examines the findings of 22 studies evaluating treatment efficacy for domestically violent males. The outcome literature of controlled quasi-experimental and experimental studies was reviewed to test the relative impact of Duluth model, cognitive-behavioral therapy (CBT), and other types of treatment on subsequent recidivism of violence. Study design and type of treatment were tested as moderators. Treatment design tended to have a small influence on effect size. There were no differences in effect sizes in comparing Duluth model vs. CBT-type interventions. Overall, effects due to treatment were in the small range, meaning that the current interventions have a minimal impact on reducing recidivism beyond the effect of being arrested. Analogies to treatment for other populations are presented for comparison. Implications for policy decisions and future research are discussed.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Duluth model; Cognitive-behavioral therapy; Recidivism; Violence

1. Introduction

As an estimated 840,000 women reported assaults at the hands of an intimate in 1996 (Bureau of Justice Statistics, 1998), interventions designed to address this growing public

* Corresponding author.

E-mail address: jbabcock@uh.edu (J.C. Babcock).

health concern have focused on the perpetrators of domestic violence in hopes of deterring further assault. Prior to the 1980s, little attention was paid to domestic violence intervention (Fagan, 1989). Issues of family privacy vs. societal best interest were paramount (Zimring, 1989); domestic violence was sometimes thought best “left behind drawn curtains” (*State v. Oliver*, 1874, cited in Rosenfeld, 1992). Subsequent criminalization of domestic violence dictated whether the crime of domestic violence should entail rehabilitation or incarceration. Since then, spouse abusers have “traditionally fallen under the rehabilitative, rather than the punitive arm of the criminal justice system” (Rosenfeld, 1992, p. 207). In actuality, with the implementation of mandatory arrest policies and court-mandated counseling, batterers’ interventions became a fusion between punishment and rehabilitation.

1.1. Current standards of care

While interventions for batterers are far from standardized, standards of care of battering interventions have been evolving in the United States since the 1990s (see Austin & Dankwort, 1999, for a review). Most states target the perpetrator as solely responsible for the crime and, as such, he shall be held accountable. Most guidelines also require training of group facilitators and experience in domestic violence work, although professional degrees and licensure are generally not required. The recommended duration of intervention ranges from 12 to 52 weeks. Finally, the group intervention model is the format of choice in 90% of mandates, and individual and couples’ therapy is deemed as inappropriate in the majority of the current standards (Austin & Dankwort, 1999). For the most part, state standards have been developed independently of empirical research.

Despite declarations that arrest followed by court-ordered treatment offers “great hope and potential for breaking the destructive cycle of violence” (U.S. Attorney General’s Task Force on Family Violence, 1984, p. 48), there is little empirical evidence that treatment is effective in reducing recidivism of family violence to any meaningful degree. In his review of the earlier studies on marital violence treatment programs, Rosenfeld (1992) concluded that men who are arrested and complete treatment have only slightly lower recidivism rates than men who are arrested but refuse treatment, dropout of treatment, or remain untreated. Some have even argued that treatment programs may put women at increased risk for domestic violence, by contributing to a false sense of security among battered women whose husbands have sought treatment (Holtzworth-Munroe, Beatty, & Anglin, 1995).

Fortunately, in the past decade, several researchers have conducted well-designed studies capable of shedding some light on questions and concerns regarding the efficacy of batterers’ treatment. A small but growing body of methodologically rigorous investigations into the effectiveness of current programs now exists. The purpose of this article is to critically review the treatment outcome research on batterers’ interventions and to conduct a meta-analysis to examine the impact of (1) the treatment type and (2) the study design on the effect size attributable to treatment. Since the current community response to battering is a combination of legal sanctions plus rehabilitation, the goal of this meta-analysis is to examine the effect of the therapeutic intervention, over and above the effect of legal interventions.

A number of studies have summarized the effects of batterers' treatment (Babcock & LaTaillade, 2000; Davis & Taylor, 1999; Hamberger & Hastings, 1993; Levesque & Gelles, 1998; Rosenfeld, 1992; Tolman & Bennett, 1990). After their review of the research literature, Hamberger and Hastings (1993, p. 220) asked the question, "What do we know about the short- and long-term effects of treatment on wife assault?" They conclude "Not much," due to methodological problems of the existing research. In his quantitative review, Rosenfeld (1992) concluded that there are minimal decreases in recidivism rates between treatment completers (36%) and men only receiving legal-system interventions (39%). Rosenfeld stopped short of conducting a meta-analysis, due to the limited number of studies using consistent methodologies available at that time. Davis and Taylor (1999) recently reviewed the empirical batterers' treatment outcome literature and came to quite different conclusions. Although they did not conduct a meta-analysis, they calculated the average effect sizes from five studies. Based on these averages, they estimated the treatment effect size to be approximately $h=0.41$ (less than 0.50 is considered "small") but nonetheless concluded that "there is fairly consistent evidence that treatment works and that the effect of treatment is substantial" (Davis & Taylor, 1999, p. 69). Levesque and Gelles (1998) were the first to presents a meta-analysis of 17 batterers' treatment outcome studies. Based on the small effect sizes (h s ranging from 0.18 to 0.27), they concluded that batterers' interventions "work a little, probably."

To help to clarify some of these discrepant conclusions, we conducted a formal meta-analysis, including the more methodologically rigorous studies, and new findings on recently completed experiments. The current study is the first formal meta-analysis on batterers' treatment outcome studies to be published to date. We attempted to improve on previous research in two ways. First, Hamberger and Hastings (1993) included studies that utilized uncontrolled, pre–post designs in their review. The level of confidence that any change in batterers' behavior was, indeed, due to treatment was undermined because extraneous causes were not ruled out by the presence of a control group. Pre–post studies preclude the estimate of an effect size due to treatment, as they are confounded with the effects of the legal system, i.e., the effects of "getting caught." As such, the present study utilized only studies that possessed some type of control group (e.g., treatment dropouts, another type of nonequivalent control group, or those randomly assigned to a no-treatment condition). Second, previous studies (Davis & Taylor, 1999; Levesque & Gelles, 1998;) have reported the effect size of batterers' treatment in terms of Cohen's h (Cohen, 1988). However, this statistic does not adjust for sample size and is more commonly used in power analysis than meta-analysis. To account for sample size, Cohen's d was selected as the measure of effect size in the present study.

1.2. Batterers' interventions

Only a few intervention modalities have been subjected to rigorous empirical test. These include feminist psychoeducational men's groups, cognitive–behavioral men's groups, anger management (a form of cognitive–behavioral group treatment), and couples' therapy.

1.2.1. Psychoeducational model

The most prominent type of clinical intervention with batterers is a feminist psychoeducational approach (Pence & Paymar, 1993). This intervention, originated by the Duluth Domestic Abuse Intervention Project program in Minnesota, is frequently referred to as the Duluth model. According to this model, the primary cause of domestic violence is patriarchal ideology and the implicit or explicit societal sanctioning of men's use of power and control over women. This program, developed from a social work perspective, typically eschews DSM-type diagnoses and does not consider the intervention to be therapy. Rather, group facilitators lead consciousness-raising exercises to challenge the man's perceived right to control or dominate his partner. A fundamental tool of the Duluth model is the "Power and Control Wheel," which illustrates that violence is part of a pattern of behavior including intimidation, male privilege, isolation, emotional, and economic abuse, rather than isolated incidents of abuse or cyclical explosions of pent-up anger or painful feelings (Pence & Paymar, 1993). The treatment goals of the Duluth model are to help men change from using the behaviors on the Power and Control Wheel, which result in authoritarian and destructive relationships, to using the behaviors on the "Equality Wheel," which form the basis for egalitarian relationships (Pence & Paymar, 1993). The feminist Duluth-type model remains the unchallenged treatment of choice for most communities. In fact, the states of Iowa and Florida mandate that battering intervention programs adhere to the general tenets of the Duluth model to be state certified (Abel, *in press*; Healey, Smith, & O'Sullivan, 1998).

1.2.2. Cognitive behavioral groups

An alternative to the feminist psychoeducational group is the cognitive-behavioral therapy (CBT) model. Cognitive behavioral batterers interventions, developed primarily by psychologists, tend to make violence the primary focus of treatment. Since violence is a learned behavior, nonviolence can similarly be learned according to the cognitive-behavioral model (Adams, 1988). Violence continues because it is functional for the user, reducing bodily tension, achieving victim compliance, putting a temporary end to an uncomfortable situation, and giving the abuser a sense of power and control (Sonkin, Martin, & Walker, 1985). Recognizing the functional aspects of violence, the cognitive-behavioral therapist points out the pros and cons of violence. In addition, they use skills training (e.g., communication, assertiveness, and social skills training) and anger management techniques (e.g., timeouts, relaxation training, and changing negative attributions) to promote awareness of alternatives to violence.

The intervention labels are often misleading. Some CBT groups are not strictly "cognitive" or "behavioral," as they address emotional components of violence, such as empathy and jealousy (Dunford, 2000). Most modern cognitive-behavior groups also usually address perpetrator attitudes and values regarding women and the use of violence toward women. To the extent that CBT groups address patriarchal attitudes, and Duluth model groups address the learned and reinforced aspects of violence, any distinction between CBT and Duluth model groups becomes increasingly unclear.

1.2.3. *Other modes of therapy*

The rationale for the use of group therapy is that men learn to confront one another's denial and victim blaming (Murphy & Baxter, 1997). As such, there have been no controlled, empirical studies to date testing individual therapy approaches for abusers. Due to concerns about the effectiveness of male-only group interventions, some in the domestic violence field are exploring alternatives to the psychoeducational group approach by testing conjoint groups (Dunford, 2000; O'Leary, Heyman, & Neidig, 1999). Advocates of couples groups state that including the wife in the group intervention may change the tenor of the men's group by rendering role-play more realistic and by reducing "women bashing" (Dunford, 2000). It may also empower the wife by allowing her to "witness authority figures confronting the offensive and oppressive nature of spouse abuse," as well as model for her constructive ways to deal with conflict (Dunford, 2000, p. 469). However, most states set standards, guidelines, or mandates that discourage or prohibit the funding of any program that offers couples or family counseling as a primary mode of intervention (Healy et al., 1998; Lipchick, Sirles, & Kubicki, 1997), as the woman's disclosures in the presence of her partner may lead to later retribution (Lipchick et al., 1997) or imply that she is at least partially to blame for her victimization (Jacobson, 1993).

2. Method

2.1. *Overview of methods of prior studies*

The primary purpose of this article is to quantitatively summarize the findings to date on the effect of batterers' treatment on violence recidivism. A review of the batterers' treatment literature was conducted using PsycInfo, entering the keywords "batterers" and "domestic violence." These were cross-referenced with terms including "treatment" and "intervention." Studies identified in this way were retrieved and their reference sections reviewed for additional treatment outcome studies. Additionally, the reference sections of five reviews of the batterer treatment literature were examined (Bargarozzi & Giddings, 1983; Davis & Taylor, 1999; Gelles, 1980; Rosenfeld, 1992; Tolman & Bennett, 1990). Prior to results of our quantitative meta-analysis, we will briefly summarize the methods and findings of available studies to the present, casting a broad net to include published materials, manuscripts in press, and data presented at national conferences. For three recent studies (Feder & Forde, 1999; Gondolf, 2000; Taft, Murphy, Elliott, & Morrel, 2001), additional information needed to calculate effect size was obtained directly from the authors.

2.1.1. *Quasi-experimental studies*

Table 1 presents the quasi-experimental studies, most of which used the nonequivalent control group design (Campbell & Stanley, 1963) to compare either treatment completers to treatment dropouts or treated offenders to a matched group of nontreated batterers (not using true random assignment). It should be noted that the nonequivalent control group design employed by most studies on battering interventions does not meet the American Psycho-

Table 1
Quasi-experimental designs

Study author	Group design and initial sample size	Treatment type	Treatment length	Attrition rates	Follow-up recidivism measure and response rates	% Re-offended	Effect size (<i>d</i>)
Taft et al. (2001), Morrel, Elliott, Murphy, and Taft (2003), and Murphy (personal communication)	Tx1 completers (<i>n</i> = 33); Tx2 completers (<i>n</i> = 41), dropouts (<i>n</i> = 12)	Tx1 = supportive + treatment retention; Tx2 = CBT + treatment retention	16 sessions	18% completed < 12 weeks	Police records at 22–36 months (73% of sample) and partner report (61% of the sample) at 6 months follow-up	Police report: Tx1 = 9.5%; Tx2 = 9.7%, dropouts = 54% Partner report: Tx1 = 10%; Tx2 = 18.5%, dropouts = 33%	Police report: Tx1 = 1.15; Tx2 = 1.22 Partner report: Tx1 = 0.69; Tx2 = 0.36
Gondolf (1997, 1998, 2000, personal communication)	Tx1 completers (<i>n</i> = 158); Tx1 dropouts (<i>n</i> = 55); Tx2 completers (<i>n</i> = 145); Tx2 dropouts (<i>n</i> = 64); Tx3 completers (<i>n</i> = 140); Tx3 dropouts(<i>n</i> = 75); Tx4 completers (<i>n</i> = 135); Tx4 dropouts(<i>n</i> = 72)	Four Duluth model programs of different lengths	Tx1 (Pittsburgh): 12 weeks with few additional services; Tx2 (Denver): 26 weeks; Tx3 (Houston): 24 weeks; Tx4 (Dallas): 12 weeks with several additional services	32% across all sites attended less than 2 months	Police reports (57%) at 15 months follow-up and cumulative partner, perpetrator, police report (48% of sample) at 30 months	Police report: Tx1 = 17%; Tx1 dropouts = 41%; Tx2 = 26%; Tx2 dropouts = 51%; Tx3 = NA; Tx3 dropouts = NA; Tx4 = 12%; Tx4 dropouts = 19% Partner report: Tx1 = 40%; Tx1 dropouts = 50%; Tx2 = 35%; Tx2 dropouts = 55%; Tx3 = 35%; Tx3 dropouts = 59%; Tx4 = 33%; Tx4 dropouts = 58%	Police report: Tx1 = 0.58 ^a ; Tx2 = 0.54; Tx3 = NA; Tx4 = 0.20 Partner report: Tx1: 0.20; Tx2: 0.41; Tx3: 0.50; Tx4: 0.52
Babcock and Steiner (1999)	Tx completers (<i>n</i> = 106); Tx dropouts (<i>n</i> = 178); incarcerated (<i>n</i> = 55)	Multisite, majority Duluth model, psychoeducational + probation	36 weeks	68% completed < 28 sessions	Police report at 2 years postprosecution	Completers = 8%; dropouts = 23%; incarcerated = 62%	Tx vs. dropouts = .40

Murphy et al. (1998)	Tx completers ($n=10$); noncompleters ($n=225$)	Duluth model psychoeducational	22 sessions	84% (of 62 men ordered to treatment) completed < full 22 weeks	Police records 12–18 months postprosecution	Completers = 0%; noncompleters = 16%	0.44
Dutton et al. (1997)	Tx completers ($n=156$); Tx dropouts and rejected ($n=290$)	Clinical anger management vs. dropouts and rejected (for noncooperation, psychosis, etc.)	16 weeks	52%	Police reports ranging up to 11 years (mean 5.2 years)	Completers = 18%; dropouts = 21%	0.07
Dobash et al. (1996)	Tx completers ($n=40$); Tx dropouts ($n=80$)	Psychoeducational group vs. dropouts	unknown	66%	Police and partner report (25% of sample) at 1 year follow-up	Police report: completers = 7%; dropouts = 10% Partner report: dropouts = 75%	Police report = 0.11 Partner report = 0.92
Newell (1994)	Tx1 = DV group completers ($n=155$); Tx1 dropouts ($n=118$); Tx2 = other Tx ($n=83$); no Tx ($n=135$)	Feminist psychoeducational group vs. other Tx (AA, couples, individual) vs. group dropouts vs. no Tx control	12 weeks	57%	Police reports (re-arrest) at 2 year follow-up	Tx1 completers = 23%; Tx1 dropouts = 36%; Tx2 = 16%; no Tx = 22%	Tx1 completers vs. dropouts = 0.29; Tx1 completers vs. no Tx = -0.02 Tx2 vs. no Tx = 0.15
Flournoy (1993)	Tx1 ($n=16$); Tx2 ($n=13$); waitlist control ($n=14$)	Tx1 = CBT; Tx2 = psychoeducational; control = waitlist	8 weeks	CBT 19%; psychoeducational 38%	Police reports 2–3 months follow-up (81% of sample)	Tx1 = 8%; Tx2 = 0%; control = 7%	Tx1 = -0.03; Tx2 = 0.33
Harrell (1991)	Tx1 ($n=81$); no-treatment control ($n=112$)	Mandated CBT group (8–12 weeks) vs. no treatment mandated	8–12 weeks	20%	Police reports at 15–29 months; partner report on 90% of sample at 6 months	Police report: Tx = 50%; no Tx = 30% Partner report: Tx = 43%; no Tx = 12%	Police report = -0.42 Partner report = -0.76
Chen, Bersani, Myers, and Denton (1989)	Mandated to Tx ($n=120$); not mandated ($n=101$)	Anger management	8 weeks	37% completed less than 7 sessions	Police reports	Completers = 5%; dropouts = 10%	0.19
Edleson and Grusznski (1988) Study 3	Tx completers ($n=84$); Tx dropouts ($n=37$)	Psychoeducation followed by process oriented	8 weeks psychoeducation + 16 weeks	31%	Partner report at 6 months follow-up	Completers = 42%; dropouts = 49%	0.14

(continued on next page)

Table 1 (continued)

Study author	Group design and initial sample size	Treatment type	Treatment length	Attrition rates	Follow-up recidivism measure and response rates	% Re-offended	Effect size (<i>d</i>)
Edleson and Grusznski (1988) Study 1	Tx completers (<i>n</i> = 27); Tx dropouts (<i>n</i> = 30)	Psychoeducation followed by process oriented	8 weeks	47%	Partner report at 6 months follow-up	Completers = 33%; dropouts = 46%	0.26
Hamberger and Hastings (1988)	Tx completers (<i>n</i> = 32); Tx dropouts (<i>n</i> = 36)	CBT group	15 weeks	53%	Combination of self + partner + police report at 1 year follow-up	Completers = 9%; dropouts = 17%	0.23
Waldo (1988)	Tx completers (<i>n</i> = 30); Tx dropouts (<i>n</i> = 30); control (<i>n</i> = 30)	Relationship enhancement men's group	12 weeks	50%	Police reports at 1 year follow-up	Completers = 0%; dropouts = 20%; controls = 20%	Completers vs. dropouts = 0.70; completers vs. control = 0.70
Leong, Coates, and Hoskins (1987)	Tx completers (<i>n</i> ≈ 33); Tx dropouts (<i>n</i> ≈ 34)	CBT group	unknown	≈ 50%	Police report at 6 months follow-up	Completers = 13%; dropouts = 29%	
Hawkins and Beauvais (1985)	Tx completers (<i>n</i> = 52); Tx dropouts (<i>n</i> = 43)	CBT	1–6 group + 6 couple and individual	45%	Police report at 6 months follow-up	Completers = 18%; dropouts = 18%	0.00
Stacey and Shupe (1984)	Initial <i>N</i> = 193; Tx1 (<i>n</i> at follow-up = 77); dropouts (<i>n</i> at follow-up = 30)	Multisite: 2 sites CBT, 1 site psychodynamic/ Rogerian	10–18 weeks	Unknown	Partner report at 0–24 month follow-up (55% of sample)	Completers = 34%; dropouts = 50%	0.33

Tx = treatment.
^a Effect sizes from the Pittsburgh site (Gondolf, 2000) were excluded from the meta-analysis due to treatment dropouts receiving additional legal sanctions.

logical Association's standards for establishing empirical support (Chambless et al., 1996). All of the quasi-experimental studies share the methodological problem of potentially "stacking the deck" in favor of treatment. Men who choose to complete treatment are known to be different from those who drop out (e.g., more educated, more likely to be employed, married, and Caucasian, and less likely to have a criminal record) (Babcock & Steiner, 1999; Hamberger & Hastings, 1988). Two studies did attempt to control for these preexisting group differences (Babcock & Steiner, 1999; Gondolf, 1997), and found that the effect attributable to treatment remained statistically significant. However, the percentages and effect sizes presented in Table 1 are not corrected for confounds due to group differences between treatment dropouts and completers. It is difficult to estimate the effect size controlling for demographic variables because most studies do not present the data in a manner such that a reanalysis, controlling for confounds, would be possible; we predict that doing so would dramatically decrease the effect size.

A second methodological difficulty with quasi-experimental designs is the degree to which the "dropout" condition is contaminated by the effects of alternative sanctions against batterers. Gondolf (2000) found that the effect size of treatment at one of his study's site (Pittsburgh) was negligible. He later learned that alternative sanctions were issued upon treatment dropouts, rendering the dropouts invalid as a comparison group. Thus, that site was excluded from our meta-analysis. Other studies that were included in the meta-analysis may have similar confounds that are undisclosed in their reports. In some cases, quasi-experimental designs and randomized experiments can yield comparable effect sizes (Heinsman & Shadish, 1996). However, whether this is true for the body of studies on batterers' intervention remains an empirical question.

2.1.2. *True experiments*

Because of the ever-present risk of confounds among quasi-experimental studies, results from randomized experiments are the "gold standard" for meta-analyses (Shadish & Ragsdale, 1996, p. 1290). Therefore, results of the five, recent experimental studies should be considered a more accurate estimate of the actual effect size due to batterers' treatment. Table 2 presents the five studies to date that have employed random assignment. These five experiments deserve special attention.

Feder and Forde (1999) randomly assigned batterers on probation to either a feminist-psychoeducational program or no treatment in Broward County, FL. In general, there were no statistically significant differences between the two groups on recidivism as measured by police records ($d=0.04$) or by victim report ($d=-0.02$). There was a small but significant effect on recidivism among the subset of men randomly assigned to group treatment who attended all 26 sessions. In this study, random assignment apparently failed, with an uneven number of men being assigned to the treatment and control condition (Feder & Forde, 1999). Moreover, this study suffered from a particularly high attrition rate of men from treatment (60%) and low response rate from victims at follow-up (22%).

In a large evaluation of U.S. Navy personnel stationed in San Diego, Dunford (1998, 2000) compared a 36-week cognitive-behavioral group and a 26-week couples therapy format to a rigorous monitoring condition and a no-treatment control (victims safety planning). Neither

Table 2
Experimental designs

Study authors	Group design and initial sample size	Treatment type	Treatment length	Attrition rates	Follow-up recidivism measure and response rates	% Re-offended	Effect size (<i>d</i>)
Feder and Forde (1999, personal communication)	Tx = Duluth (<i>n</i> = 174); control (<i>n</i> = 230)	Duluth + probation vs. probation only	26 weeks	60%	Police at 1 year and partner report (22% of sample) at 6 month follow-up	Police report: Tx = 4.8%; control = 5.7% Partner report: Tx = 32.7%; control = 31.6%	Police report: Tx = 0.04 Partner report: Tx = − 0.02
Dunford (2000)	Tx1 = CBT (<i>n</i> = 168); Tx2 = couples (<i>n</i> = 153); monitoring (<i>n</i> = 173); control (<i>n</i> = 150)	CBT men's group, conjoint Tx, and rigorous monitoring vs. victim safety planning control	Tx1 = 36 weeks + 6 monthly meetings; Tx2 = 26 weeks + 6 monthly meetings; monitoring = monthly meetings for 12 months	29%	Police and partner report (72% of initial sample of 861) on at 1 year follow-up	Police report: Tx1 = 4%; Tx2 = 3%; monitoring = 6%; control = 4% Partner report: Tx1 = 29%; Tx2 = 30%; monitoring = 27%; control = 35%	Police report: Tx1 = 0.00; Tx2 = 0.05; monitoring = − 0.09 ^a Partner report: Tx1 = 0.13; Tx2 = 0.10; monitoring = 0.17 ^a

Davis et al. (2001)	Tx1 = long ($n = 129$); Tx2 = brief ($n = 61$); control ($n = 186$)	Duluth model treatments vs. community service control	Tx1 = 26 weeks; Tx2 = 8 weeks	33%	Police and partner report of new incident in past 2 months (50% of sample) at 1 year follow-up	Police report: Tx1 = 10%; Tx2 = 25%; control = 26% Partner report: Tx1 = 14%; Tx2 = 18%; control = 22%	Police report: Tx1 = 0.41; Tx2 = 0.02 Partner report: Tx1 = 0.21; Tx2 = 0.10
Ford and Regoli (1993)	Tx1 = pretrial diversion into counseling ($n = 127$); Tx2 = counseling as condition of probation ($n = 114$); control = sentence without counseling ($n = 106$)	Counseling (unknown type) as pretrial diversion vs. condition of probation vs. other sentencing (e.g., fine, jail) control	Unknown	Unknown	Partner report at 6 month follow-up (31% of sample)	Tx1 = 34%; Tx2 = 45%; control = 34%	Tx1 = 0.00; Tx2 = - 0.22
Palmer et al. (1992)	Tx ($n = 30$); control ($n = 26$)	Psychoeducational vs. probation only	10 weeks	30% attended < 7 sessions	Police at 1–2 year follow-up	Tx = 10%; control = 31%	Tx = 0.54

Tx = treatment.

^a Effect sizes generated from the rigorous monitoring conditioning (Dunford, 2000) were excluded from this meta-analysis, as it does not represent a therapeutic intervention. Weighted percentage of nontreated who re-offended based on police report = 21%; based on partner report = 35%.

CBT men's groups ($d=0.13$) nor couples therapy ($d=0.10$) had a significant impact on recidivism at 1-year follow-up based on victims' report. This study represents the most methodologically rigorous study conducted to date in terms of sample size, length of follow-up, attrition rates, follow-up reporting rates, and assessment of treatment adherence. However, it is important to note that this sample of batterers, those employed through the Navy in San Diego, are not representative of the population of batterers court-mandated to domestic violence programs around the country. All of the research participants were employed, had a high stake in social conformity, and thus, were more "socially bonded" (Sherman, Smith, Schmidt, & Rogan, 1992). Any intervention, including arrest and being identified by authorities, may work to deter socially bonded individuals from repeat offenses. This may be reflected in the unusually low official recidivism rates of the nontreated batterers (4%).

Davis, Taylor, and Maxwell (2001) compared a long (26-week) psychoeducational group to a brief (8-week), psychoeducational group, and to a community service control (70 hours of clearing vacant lots, painting senior citizen centers, etc.) in Brooklyn, NY. They found a statistically significant reduction in recidivism and a small but respectable effect size of $d=0.41$ based on criminal records among the long treatment group only; the 8-week group was indistinguishable from the community service control ($d=0.02$). As for partner report, this study employed a rather unusual method of calculating re-offenses. Only new incidents of violence in the 2 months prior to the follow-up contact point were included rather than a cumulative count. When based on victim report of these recent offenses, neither the long nor the brief intervention had a statistically significant effect on re-assault when compared to no treatment. Correspondingly, the effect size due to treatment based on partner report of subsequent violence was small ($d=0.21$). It is important to note that, like in the Broward County experiment (Feder & Forde, 1999), random assignment may have been compromised. In the Brooklyn experiment (Davis et al., 1998), nearly 30% of initial assignments were subjected to "judicial overrides" (Gondolf, 2001); that is, judges reassigned defendants to different interventions.

Ford and Regoli (1993) designed a study that randomly assigned batterers into treatment as a pretrial diversion (i.e., defendants' criminal records would be cleared pending treatment completion), treatment as a condition of probation postconviction, vs. alternative sentencing strategies (e.g., paying a fine or going to jail). Although this study was designed to test different sentencing options rather the effects due to treatment, one can compare batterers sentenced to treatment vs. batterers not sentenced to treatment (although the type of treatment and actual attendance rates were not specified). Again, there were no significant differences or effect sizes comparing recidivism rates based on victim report between men sentenced to treatment vs. those who were not. Neither treatment as pretrial diversion ($d=0.00$) nor as a condition of probation postconviction ($d=-0.22$) was found to be superior to purely legal interventions.

Finally, Palmer, Brown, and Barrera (1992) conducted a small scale study in Canada of men using block random procedure: men were assigned to 10-week psychoeducational treatment if a new group was to commence within 3 weeks or, if not, to a "probation only" control group. The relatively unstructured, client-centered treatment addressed beliefs about

violence, responsibility for violent behavior, coping with conflict and anger, self-esteem, and relationships with women (Peluch, 1987). Based on police reports, men assigned to the treatment condition re-offended at a significantly lower rate than men assigned to probation only, yielding a medium effect size ($d=0.54$). However, this study is limited by its small sample size, and the results may not be generalizable to other samples.

Conducting an experiment in which judicial discretion is sacrificed and criminals are randomly assigned to treatment or no treatment can be problematic on ethical as well as practical grounds (Dutton, Bodnarchuk, Kropp, & Hart, 1997). Adopting an experimental design does not guarantee a more rigorous evaluation than quasi-experimental designs afford (Gondolf, 2001). While it is true that experimental designs permit greater confidence in conclusions regarding causal relations, it is also the case that problems with differential attrition and failure of random assignment reduce internal validity of this design. Additionally, researchers must grapple with the “intention-to-treat” problem: should effect sizes be calculated from the initial sample size or from the completers only? What if the majority of “treated” offenders attended no-treatment groups whatsoever? It is recommended that researchers report both recidivism rates for all batterers who were assigned to treatment as well as those who actually completed treatment (although few of studies have done so).

2.2. *Study inclusion criteria*

Originally, 68 empirical studies of the efficacy of batterers’ treatment programs were located. These studies were classified according the design: experimental ($k=5$), quasi-experimental ($k=17$) and pre–post ($k=48$). The criterion for inclusion in this meta-analysis was the (1) the presence of some form of comparison group of batterers and (2) reliance on victim report or police record as the index of recidivism. The uncontrolled, pre–post test studies have been reviewed previously (Davis & Taylor, 1999; Hamberger & Hastings, 1993; Rosenfeld, 1992). These are the weakest methodological designs and generally tend to overestimate effect size (Lipsey & Wilson, 1993). On this basis, 48 of the 70 studies were not included. The stronger quantitative evaluations of domestic violence interventions generally fall into two categories: (1) quasi-experimental, where treatment completers are compared to treatment dropouts or to a matched comparison group that did not receive treatment and (2) true experimental designs, where clients are randomly assigned to treatment(s) vs. no treatment. Studies ($k=22$) consisting of experimental and quasi-experimental designs formed the data for this quantitative review. These studies yielded at total of 44 effect sizes (effect sizes formed the unit of analysis for the present study), in which a treatment group was compared to either a randomized control or treatment dropouts.

Several recent studies have compared two active treatments for domestic violence without the inclusion of a control group and without the comparing of completers to a no-treatment or dropout comparison group. For example, two studies (Brannen & Rubin, 1996; O’Leary et al., 1999) compared a couples format to a gender-specific group format. Saunders (1996) compared the response of batterers with different personality profiles to both more standard structured groups vs. more process-oriented therapies. Although one could calculate an effect size from these treatment comparison studies, the resultant statistic would reflect the

magnitude of the difference between two active treatments. This would grossly underestimate the effect size due to treatment and potentially bias the results against treatment. Other studies (e.g., Edleson & Syers, 1991) report recidivism rates for the treatment completers only. Sullivan and Bybee (1999) conducted a well-designed randomized clinical trial of victims' advocacy intervention that yielded an effect size (d) of 0.35 based on victim report of recidivism; however, interventions with victims are beyond the scope of the current meta-analysis. While these studies are important contributions to the clinical literature, they do not render effect sizes estimating the effect due to batterers' treatment. As such, they are not included in this meta-analysis.

2.3. Coding

Effect sizes were coded along a number of dimensions that were theoretically promising for investigation as moderators. Each effect size was classified according to the type of report upon which recidivism rates were based, treatment type, treatment duration, follow-up time, and attrition rates. Report of batterer recidivism took two forms: police report and partner report. Many of the earlier effect sizes relied exclusively on batterers' self-report as an outcome measure. Such effects cannot differentiate between treatment success and batterers' tendency to vastly underreport the true incidence of abuse (Davis & Taylor, 1999; Rosenfeld, 1992). Moffit et al. (1997) found that the reliabilities between male and female report in a community sample on the presence/absence of violence was poor (average kappa=.36). They concluded that in a therapeutic or correctional setting, "where the pressures (for batterers) to bias their reports may outweigh researchers' promises of confidentiality," collateral reports may be essential (Moffit et al., 1997, p. 54). In light of this potential reporting bias, only effect sizes that use at least one independent report of recidivism, either victim report or criminal record, were included in this review. In many cases, both police and partner reports were examined. As such, our analyses of the data examined separate effect sizes for all comparisons presented in each study; moderator analysis was performed twice, separately for partner and police report, to avoid artificial inflation of the number of studies.

2.3.1. Recidivism

Considering practical significance, most clinicians working with batterers agree that cessation of intimate partner violence is an important success criterion (Edleson, 1996), rather than, for example, showing a decrease in the frequency or severity of violent acts. For the purposes of this review, "recidivism" is considered any report of physical violence reported by the victims and/or any domestic violence incidents reported to the police during a follow-up period (i.e., recidivism is a dichotomous variable and the effect sizes are calculated from the proportion of men who re-offended). Most studies utilized the Conflict Tactics Scale (CTS/CTS-2) (Straus, 1979; Straus, Hamby, Boney-McCoy, & Sugarman, 1996) as an outcome measure of partner report of recidivism. Our dependent measure of partner report of recidivism was the percentage that stopped further physical aggression, rather than change in CTS score. While a cessation of men's emotional abuse and increased accountability are

also goals of many treatment programs, few program evaluations or experiments have examined any outcome variable other than physical abuse frequency or criminal recidivism. Therefore, only police report and partner reports of physical assault were included in this meta-analytic review.

2.3.2. *Treatment types*

Treatment types included Duluth/feminist psychoeducational ($k=19$), CBT ($k=11$), and other ($k=7$). Since the meta-analytic method requires examination of the heterogeneity/homogeneity of variance due to each putative moderator, a cell size greater than one was required for each level of the moderator. As a consequence, any type of treatment that occurred only once in the database was aggregated into the “other” category. The seven effect sizes in the “other” category came from studies testing the effectiveness of couples therapy (Dunford, 2000), supportive therapy (Taft et al., 2001), relationship enhancement (Waldo, 1988), a mixture of different interventions (Newell, 1994), and therapies of an unspecified type (Ford & Regoli, 1993).

2.3.3. *Treatment length*

Treatment length was dichotomized: short (mean treatment length <16 weeks), and long (mean treatment length 16 weeks). If any treatment did not maintain a uniform duration, the average length of treatment was utilized.

2.3.4. *Follow-up length*

Effect sizes were classified into one of two categories based on follow-up length: short (mean follow-up time <12 months) and long (mean follow-up time >12 months). For studies with variable follow-up times, the mean follow-up time was calculated.

2.3.5. *Attrition*

Attrition from treatment was calculated as the percentage of individuals who were classified as “dropouts” from the quasi-experimental studies by the authors. It should be noted that different authors have distinct criteria for what constitutes treatment completion. For some effect sizes, completers must attend 100% of the sessions; other authors report “completers” as those attending 80% or more of the required sessions. Due to the inconsistencies in calculating and reporting attrition, this variable was not entered into the meta-analysis. However, attrition rates are reported in Tables 1 and 2. They may be viewed as an index of quality of treatment or quality of the coordinated community response and may influence the effect size.

2.4. *Reliability*

For reliability purposes, both the first and second authors reviewed and coded each study. There were no disagreements on study design, type of report, length of treatment, or follow-up length (reliability = 100% agreement). However, there was one study in which the coders disagreed on the treatment type (reliability = 95% agreement). In this case, the study author

was contacted (Jeffrey Edelson, personal communication, September 13, 2000) to assist in assigning a label to the treatment.

2.5. Estimates of effect size

Table 1 presents the general design, type of treatment, and recidivism or re-offense rates of all identifiable quasi-experimental designs, and Table 2 the existing true experimental studies conducted in the past decade. The re-offense rates (that is, the percentage in the treated and control conditions who re-offended) as reported in the studies were then recalculated into an effect size, using the g statistic on proportions (Hedges & Olkin, 1985). The g statistic on proportions was then transformed into the d statistic, adjusting for sample size (Johnson, 1995). It is important to note that the size of the final samples with complete recidivism data at follow-up, especially those based on partner/victim report, is usually significantly smaller than the initial n . In many cases, the specific n s of treated and comparison groups with complete follow-up data were not explicit, although the follow-up response rate usually was. In those cases, we estimated the final n by “discounting” the initial n in each condition by the proportion with complete follow-up data.

An “effect size” is an attempt to quantify the magnitude of the effect due to treatment using a shared metric that is not influenced by the size of the sample. When based on the d statistic, effect sizes of 0.20 are considered “small,” 0.50 are considered “medium,” and effect sizes 0.80 and above are considered large (Cohen, 1988). The d effect size is in units of standard deviations; therefore, an effect size of 0.25 translates to an improvement of one-fourth of a standard deviation compared to no treatment. In true experimental designs, the effect size allows us to evaluate the magnitude of the impact that treatment has on recidivism; in quasi-experimental designs, the effect size approximates the strength of relationship between treatment and recidivism, uncontrolled for external confounds (Campbell & Stanley, 1963).

Effect sizes and variances were calculated in terms of d using Hedges and Olkin’s (1985) meta-analytic method. This enabled differential weighting of effects for sample size. Calculation of the d was accomplished utilizing D-Stat version 1.11 (Johnson, 1995). This software program calculates d based on proportions by treating each proportion as the mean of a distribution of successes vs. failures. Effect sizes were computed for each comparison for each dependent measure (i.e., report type), resulting in a total of 37 effect sizes. Moderator analysis was then conducted using MetaWin 1.0 (Rosenberg, Adams, & Gurevitch, 1997). This computer program follows Hedges and Olkin’s hierarchical approach to meta-analysis that employs the Q statistic to determine the degree of heterogeneity that exists between and within groups. As mentioned previously, other studies (Davis & Taylor, 1999; Levesque & Gelles, 1998) have reported the effect size of batterers’ treatment in terms of Cohen’s h (Cohen, 1988). Recalculating the effect sizes in terms of Cohen’s h does not substantially change the conclusions of this article. The d effect sizes can easily be converted to r effect sizes (Wolf, 1986, p. 35)¹ to calculate a

¹ Formula for r -to- d transformation: $r = \frac{d}{\sqrt{d^2 + 4}}$.

binomial effect size display (BESD), using the formula (Rosenthal, 1995; Rosnow & Rosenthal, 1988):

$$\text{BESD} = 0.50 + (r/2)$$

The BESD allows for translation of the effect size in terms of differential outcome rates to assist in interpreting the practical importance of the effect size.

Previous works (Babcock & LaTaillade, 2000; Davis & Taylor, 1999) have informally examined the effect of batterers' treatment by taking the average effect size across study. In contrast, formal meta-analyses weight effect sizes by sample size. Therefore, the results of this article may differ substantially from simply averaging or “eyeballing” of the effect sizes presented in the tables.

3. Results

Based on the data summarized in Table 1, the weighted percentage of nontreated offenders who recidivated was 21% based on police reports and 35% based on partner reports. These recidivism rates for nontreated offenders are consistent with those previously reported (O’Leary et al., 1989; Rosenfeld, 1992).

3.1. Publication bias

Analysis for publication bias and the “file drawer” phenomenon was conducted using a normal-quantile plot (Wang & Bushman, 1998). If null findings were selectively ignored, the normal-quantile plot would reveal absence of effect sizes around zero. Examination of the plots revealed no evidence for a publication bias (see Fig. 1).

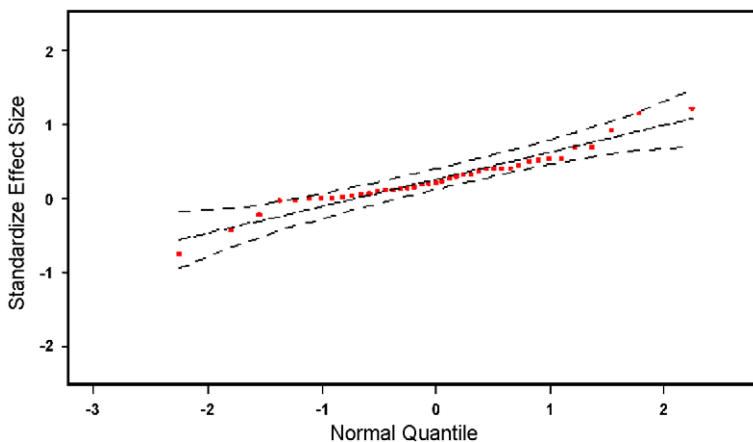


Fig. 1. Normal quantile plot to assess for the “file-drawer” problem.

3.2. Outlier analysis

Outlier analysis was conducted using the sample adjusted meta-analytic deviancy statistic (SAMD) developed by Huffcutt and Arthur (1995). The SAMD statistics were calculated separately for police and partner report. Examination of the scree plot of SAMD statistics when recidivism was assessed by police report suggested four possible outliers: both CBT (SAMD=8.73) and supportive interventions (SAMD=6.99) with retention techniques reported by Taft et al. (2001) and CBT in Harrell (1991) (SAMD=−11.08). Taft et al. and Harrell were thus excluded from subsequent analyses.

The scree plot of SAMD statistics based on partner report indicated that there were two outliers. These data points represented Dobash, Dobash, Cavanagh, and Lewis (1996) and Harrell (1991) with SAMDs of 11.01 and −15.02, respectively. Both effect sizes were excluded from the subsequent analysis based on outlier analysis.

3.3. Moderators of effect size

The remaining 36 effect sizes were entered into the hierarchical fixed effects analysis described by Hedges and Olkin (1985). A model was tested that reflected a combination of methodological and treatment moderators (Fig. 2); these included: report type, experimental vs. quasi-experimental design, and treatment type.

3.4. Effects due to method of assessing recidivism

The first moderator variable entered into the analysis was report type. The resulting analysis of two sets of effect sizes based on police and partner reports (i.e., a hierarchical moderator approach) permitted optimal use of the existing data without redundant use of samples in each group. We report 95% confidence intervals (CIs) for all effect size estimates. CIs that do not contain zero can be considered statistically significant from zero at the $P < .05$ level. Effects based on police report ($k = 20$) yielded an overall effect size of $d = 0.18$ (95% CI = 0.11–0.25) and the effects based on partner report ($k = 16$) yielded an equivalent effect size of $d = 0.18$ (95% CI = 0.08–0.28). Examination of the Q -within statistic was not significant heterogeneity for police report ($Q_w = 26.96$, $df = 19$, ns) or partner report ($Q_w = 10.96$, $df = 16$, ns). A significant Q_w statistic indicates heterogeneity among the effect sizes that suggested the existence of further moderators. While the Q_w was not statistically significant for either police or partner report, indicating a lack of heterogeneity, the presence of the hypothesized model (Fig. 2) warranted continued examination of the remaining moderators (Rosenthal, 1995).

3.5. Effect due to study design

The second moderator variable entered into the model was research design (i.e., experimental or quasi-experimental). This variable was examined for effects based on police and partner report. Analysis of research design as a moderator for effect size within police report revealed that experimental designs ($k = 6$) had an overall $d = 0.12$ (95% CI = 0.02–0.22). The

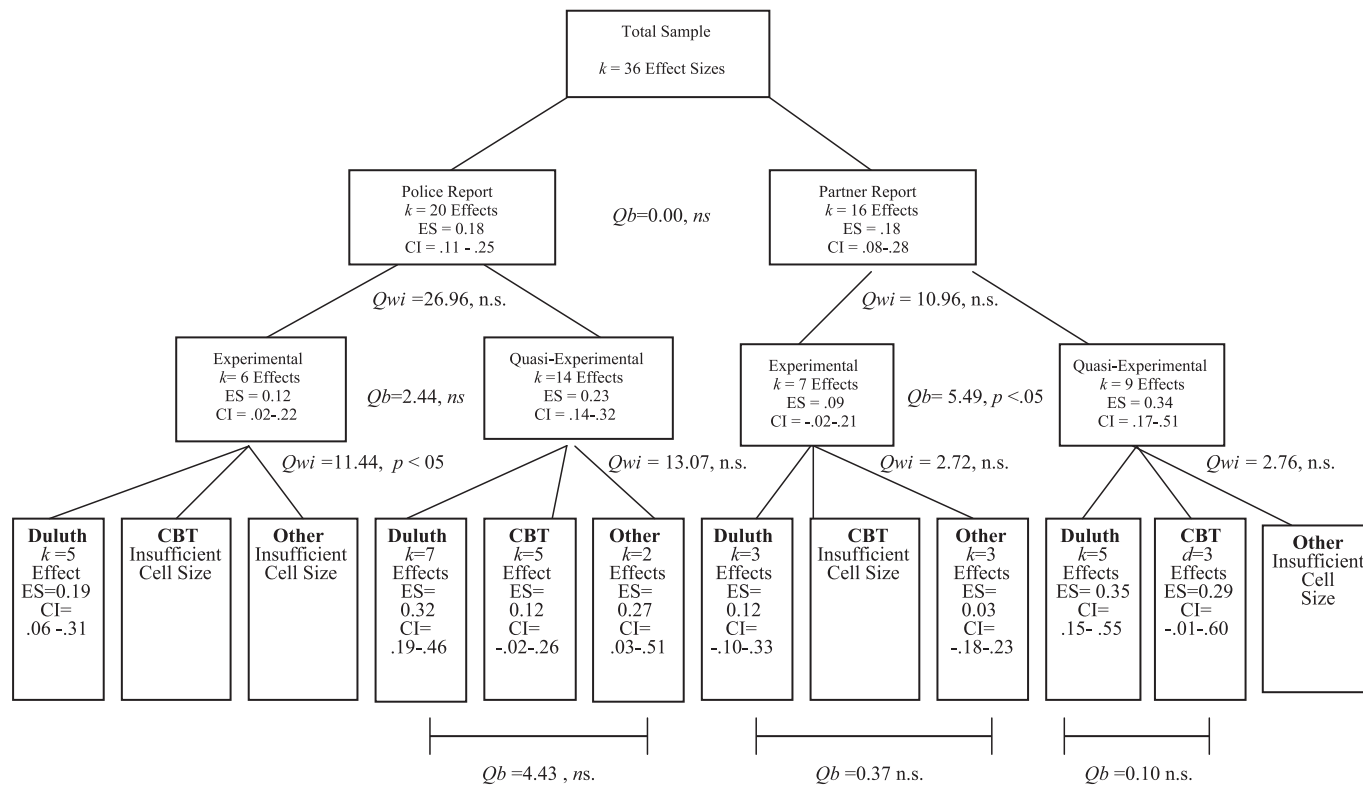


Fig. 2. Meta-analytic model testing recidivism report, study design, and type of treatment as moderators.

overall effect size for quasi-experimental designs with police report ($k = 14$) was $d = 0.23$ (95% CI = 0.14–0.32). For both experimental and quasi-experimental designs, treatment had a significant yet small impact on the cessation of domestic assault. There was not a significant difference between the overall effect sizes for experimental and quasi-experimental designs based on police report ($Q_b = 2.44$, $df = 1$, ns). Examination of results based on police report indicated that there was significant heterogeneity among effect sizes among experimental designs ($Q_w = 11.44$, $df = 5$, $P < .05$.) but not for quasi-experimental designs ($Q_w = 13.07$, $df = 13$, ns).

Similar analyses were conducted for effect sizes based on partner report ($k = 16$). Analysis of research design as a moderator for effect size within partner report revealed an average effect size for experimental designs ($k = 7$) of 0.09 (95% CI = – 0.02–0.21), not significantly different from zero. The overall effect size based quasi-experimental designs with partner report ($k = 9$) was $d = 0.34$ (95% CI = 0.17–0.51). This represents a significant yet small effect size. There was a statistically significant difference between the overall effect sizes for experimental and quasi-experimental designs based on partner report ($Q_b = 5.49$, $df = 1$, $P < .05$.). Examination of the effect sizes based on experimental designs and partner report indicated that there was not significant heterogeneity ($Q_w = 2.72$, $df = 6$, ns). Inspection of the effect sizes based on quasi-experimental designs and partner report indicates that there is not significant heterogeneity ($Q_w = 2.76$, $df = 8$, ns) within these cells.

3.6. *Effect due to treatment type*

The third moderator variable entered into the model was treatment type. This was examined for effect sizes based on experimental design and police report, quasi-experimental design and police report, and quasi-experimental design and partner report. Calculation of the effect overall effect size due to treatment type within experimental designs with police report indicated that Duluth ($k = 5$) had an effect size of $d = 0.19$ (95% CI = 0.06–0.31). CBT and “other” therapies lacked sufficient cell size ($k < 2$) to calculate an effect size. Thus, Duluth demonstrated a small effect based on police report and experimental design.

Examination of the effect overall effect size due to treatment type within quasi-experimental designs with police report indicated that Duluth ($k = 7$) had an effect size of $d = 0.32$ (95% CI = 0.19–0.46), CBT ($k = 5$) had an effect size of $d = 0.12$ (95% CI = – 0.02–0.26), and other ($k = 2$) had an effect size of $d = 0.27$ (95% CI = 0.03–0.51). In this case, the effect sizes from the Duluth model and “other” interventions were significantly different than zero, whereas CBT interventions were not significantly different from zero. However, these effect sizes did not differ significantly from one another ($Q_b = 4.43$, $df = 2$, ns).

Examination of the overall effect size due to treatment type within experimental designs with partner report indicated that Duluth ($k = 3$) had an effect size of $d = 0.12$ (95% CI = – 0.10–0.33) and other ($k = 3$) had an effect size of $d = 0.03$ (95% CI = – 0.18–0.23). CBT therapies lacked sufficient cell size ($k < 2$) to calculate an effect size. Effect sizes did not differ significantly from one another ($Q_b = 0.37$, $df = 2$, ns).

Calculation of the overall effect size due to treatment type within quasi-experimental designs with partner report indicated that Duluth interventions ($k = 5$) had an effect size of $d = 0.35$ (95%

Table 3

Effect size due to factors not tested in the model

	<i>d</i> based on police report	<i>k</i>	95% CI	<i>d</i> based on partner report	<i>k</i>	95% CI
Overall treatment length	0.18	18	0.11–0.25	0.20	14	0.10–0.32
Long	0.16	8	0.08–0.25	0.18	8	0.06–0.31
Short	0.20	10	0.09–0.32	0.30	6	0.08–0.51
Overall follow-up length	0.18	19	0.11–0.25	0.18	16	0.08–0.28
Long	0.25	8	0.14–0.35	0.48	3	0.21–0.75
Short	0.13	11	0.04–0.22	0.13	13	0.02–0.24
Overall attrition rate	0.18	19	0.11–0.25	0.20	14	0.09–0.31
High ($\geq 50\%$)	0.20	8	0.10–0.30	0.09	2	–0.23–0.40
Low ($< 50\%$)	0.16	11	0.07–0.26	0.22	11	0.10–0.34

k = number of studies.

CI = 0.15–0.55) and CBT ($k = 3$) had an effect size of $d = 0.29$ (95% CI = -0.01 – 0.60), while the “other” category lacked sufficient cell size to be included in this analysis. The two effect sizes, however, were not significantly different from each other ($Q_b = 0.10$, $df = 1$, ns).

Due to the small cell sizes for each treatment type a second model was examined that aggregated the experimental and quasi-experimental effect sizes for each reporting method (i.e., police or partner). Q -within and Q -between statistics were identical to the initial model for police and partner report. Calculation of the overall effect size due to treatment type within police report indicated that Duluth ($k = 11$) had an effect size of $d = 0.25$ (95% CI = 0.16 – 0.34), CBT ($k = 6$) had an effect size of $d = 0.09$ (95% CI = -0.03 – 0.20), and other ($k = 3$) had an effect size of $d = 0.09$ (95% CI = -0.01 – 0.32). There were no statistically significant differences between effect sizes for among the three treatment categories ($Q_b = 4.80$, $df = 2$, ns).

Examination of the overall effect size due to treatment type within partner report indicated that Duluth ($k = 8$) had an effect size of $d = 0.24$ (95% CI = 0.09 – 0.39), CBT ($k = 4$) had an effect size of $d = 0.20$ (95% CI = -0.001 – 0.40), and other ($k = 4$) had an effect size of $d = 0.04$ (95% CI = -0.16 – 0.25). There were no statistically significant differences between effect sizes for among the three treatment categories ($Q_b = 2.36$, $df = 2$, ns).

There was inadequate power to assess effect due to treatment length or follow-up length as moderator variables under different types of treatment. The overall effect sizes for treatment length, follow length, and attrition are reported in Table 3. Further analysis was conducted to examine the degree to which the inclusion of outliers in the analysis altered the present findings. In particular, a “best-case” scenario was evaluated in which only the low outliers were excluded from the analysis. The results were not significantly different from the model with all outliers removed.

4. Discussion

In general, the effect size due to group battering intervention on recidivism of domestic violence is in the “small” range. There were no significant differences in average effect size

between Duluth-type and cognitive–behavioral battering intervention programs using either police records or victim reports as the index of recidivism. While quasi-experimental designs tended to yield higher effect sizes than true experiments, the differences in effect sizes were not significant. Regardless of reporting method, study design, and type of treatment, the effect on recidivism rates remains in the small range. In the best case scenario, using quasi-experimental designs based on partner report, the effect size is $d=0.34$ indicating that treated offenders showed a one-third standard deviation in improvement in recidivism as compared to nontreated batterers.

If one relies exclusively on the five experimental studies, the effect sizes are even smaller. However, the effect sizes may be small as a result of measurement error and methodological difficulties common to research in applied settings (McCartney & Rosenthal, 2000). McCartney and Rosenthal (2000, p. 178) warn that “(g)iven that the stakes are so high, we should be wary of accepting the null hypothesis when it might very well be false—as it almost always is.” Based on the experimental studies, the effect size (d) due to treatment is 0.09 and 0.12, based on victim report and police records, respectively. This means that treatment is responsible for an approximately one-tenth of a standard deviation improvement in recidivism. Based on a partner report, treated batterers have a 40% chance of being successfully nonviolent, and without treatment, men have a 35% chance of maintaining nonviolence. Thus, there is a 5% increase in success rate attributable to treatment. To a clinician, this means that a woman is 5% less likely to be re-assaulted by a man who was arrested, sanctioned, and went to a batterers’ program than by a man who was simply arrested and sanctioned. Whether this success rate is cause for celebration or despair depends on a cost–benefit analysis; taking into account the cost of treatment and any potential “side effects” vs. the benefits of injury prevention and decreased psychological risk to the victim as well as the children exposed to family violence. A 5% decrease in violence may appear insignificant; however, batterers treatment in all reported cases of domestic violence in the United States would equate to approximately 42,000 women per year no longer being battered.

4.1. How large of an effect size should we expect?

One way to contextualize the effect size due to treatment is by comparing it to the effect sizes for treatment in other populations. Davis and Taylor (1999) compared their treatment effect size of 0.41 to the effect size of an early clinical trial on the effect of aspirin on heart attacks, which was only 0.068 and constitutes a 4% reduction in heart attacks (Rosnow & Rosenthal, 1988). Compared to this standard, they conclude that “the effect sizes seen in batterers’ treatment studies are quite substantial” (Davis & Taylor, 1999, p. 85). However, the average effect size across psychotherapy studies is much larger, approximately $d=0.85$ (Smith, Glass, & Miller, 1980). In practical terms, psychotherapy leads to benefits in 70% of cases (Rosenthal, 1995). Compared to this standard, there is great room for improvement in our batterers’ treatment interventions.

However, comparison with psychotherapy outcomes in general may not be fair. Most psychotherapies address internalizing problems (e.g., depression, anxiety) rather than

externalizing problem behavior, like aggression. Given that aggression is difficult to treat, compounded with the fact that batterers are generally not seeking treatment voluntarily and do not necessarily expect the interventions to help (Gondolf, 2001), perhaps an overall small effect due to treatment is to be anticipated. A recent meta-analysis of psychotherapy with children and adolescents reveals that the effect size for treatments of aggression was $d=0.32$ (Weisz, Weiss, Han, Granger, & Morton, 1995), indicating a 16% improvement in success rate over no treatment. Correctional treatments with adult prisoners result in effect sizes averaging $d=0.25$ (Loesel & Koefel, 1987, cited in Lipsey & Wilson, 1993), approximating a 12% improvement rate. Based on Rosenfeld's (1992) earlier review of the literature, Dutton (1998, p. 177) speculated that the effects of battering interventions fall midrange between the effects due to psychotherapy and the effects due to rehabilitation of offenders. Results from this meta-analysis reveal that even Dutton's rather modest claim appears to be overly optimistic. The effects due to battering intervention are much closer to rehabilitation effects than the effect sizes of psychotherapy in general.

4.2. Have "all won and all must have prizes?"

While the effect attributable to treatment is to some extent dependent on the methodologies employed by the studies, the effect sizes for Duluth model and CBT treatments remain relatively similar. With liberal estimates based on quasi-experimental studies, Duluth interventions yield a small effect size of $d=0.35$ while CBT interventions yield a smaller effect size of $d=0.29$. Given the variability in effect sizes of the studies that make up these averages, however, we cannot say that CBT is outperformed by Duluth-type treatment. While some may attempt to selectively use these data to bolster their arguments, claims for the superiority of one treatment type over another is unwarranted.

In retrospect, it is not surprising that there were no significant differences between CBT and Duluth-type interventions. Modern batterer groups tend to mix different theoretical approaches to treatment, combining both feminist theory of power and control as well as specific interventions to deal with anger control, stress management, and improved communication skills (Davis & Taylor, 1999; Healy et al., 1998). The "brand name" labels can be misleading. No researchers to date have conducted a head-to-head comparison between CBT and Duluth-type battering interventions, perhaps due the difficulty in identifying treatment techniques unique to either school.

It is common in the psychotherapy outcome literature to find that different modalities of treatment are equally effective—and to conclude that all have won (Beutler, 1991). This phenomenon of finding comparability in treatment outcomes is referred to as the "dodo bird verdict" (Beutler, 1991; Luborsky et al., 1975). Equivalent effect sizes due to treatment are common results of comparative studies of two active treatments (DeRubeis & Crits-Cristoph, 1998). In this case, only one study has conducted a randomized clinical trial of two active treatments (CBT and couples groups) against a no-treatment control (Dunford, 2000). Within this study and across the domain of studies to date, effects sizes due to all types of interventions are small.

4.3. Have all lost?

While the effect size due to treatment overall is in the small range, there are some specific studies finding large effect sizes. As shown in [Table 1](#), the interventions with the largest effect sizes were obtained from 16-week group therapies supplemented with retention techniques ([Taft et al., 2001](#)) and 12-week relationship enhancement skills training groups ([Waldo, 1988](#)). These findings can either be dismissed as “outliers” among scientific treatment studies, or viewed as harbingers of potentially powerful interventions. In the first study, [Taft et al. \(2001\)](#) randomly assigned men to either CBT or supportive therapy groups, both of which were supplemented with techniques designed to improve treatment retention based on the principles of motivational interviewing ([Miller & Rollnick, 1991](#)). These techniques consisted of reminder phone calls and supportive handwritten notes after intake and after missed sessions. As a result, the authors report one of the lowest attrition rates in the literature. The core therapies differed dramatically from one another, one being highly structured and the other unstructured, but both revealed strong effect sizes, especially when based on police report. This study suggests that the small effect sizes due to batterers’ interventions may be in part attributable to the client’s noninvestment and subsequent attrition from the programs. These simple techniques, which can be an adjunct to any type of program, may increase the client’s perception that the program is aware of his absence and is invested in his welfare. Thus, he may be more motivated to complete and actively participate in the program, lowering attrition and recidivism.

The second study to find a large effect size was an evaluation of an intervention called relationship enhancement ([Guerney, 1977](#)). The goals of relationship enhancement as applied to battering are to help the men develop interpersonal skills that enhance relationships and enable them to stop their use of violence ([Waldo, 1988](#)). Interventions include role-plays and assigned homework targeted to improve expressive skills, empathy, communication with the partner, and the identification and management of their emotions (see [Waldo, 1985](#)). This study suggests that more emotion-focused, rather than cognitively focused, interventions may increase the effect size of batterers treatment. Of course, the results of any single, unreplicated study should not be over generalized. More research is needed on the effectiveness of motivational interviewing as well as emotion-focused approaches as treatment modalities or as additive components to existing batterers’ intervention groups.

4.4. Limitations

One of the greatest concerns when conducting a meta-analysis is the ease at which the “bottomline” is recalled and the extensive caveats for caution are forgotten or ignored. Although we selected only studies that met our minimal criteria for rigor (inclusion of a comparison group, a follow-up period beyond the end of treatment, not relying on batterers’ self-report), there remains significant variability in the quality of research studies. Even the experimental studies are hindered by problems with high attrition rates, inconsistencies in reporting recidivism for dropouts, and low reporting rates at follow-up ([Gondolf, 2001](#)). Some of these factors that affect the quality of the research studies are confounded with

treatment quality and quality of the community response, broader factors that cannot always be ascertained. Therefore, caution in interpreting these results is warranted. Meta-analyses are only as robust as the individual studies taken into account.

Quasi-experiments make up the bulk of the studies included in this meta-analysis, but studies comparing treatment completers to dropouts are inherently confounded by self-selection. Quasi-experiments capitalize on “creaming” (Davis & Taylor, 1999); that is, comparing the most highly motivated batterers with the least motivated batterers, “thereby stacking the deck in favor of finding program effects” (Davis & Taylor, 1999, p. 74). Yet, experiments have interpretational biases as well. Most studies base outcomes according to the original random assignment. If the experimental treatment suffers from high attrition rates, and the outcome data are based on “intention to treat,” there is a strong possibility that few people received an adequate “dose” of treatment (Gondolf, 2001). The alternative, using treatment actually received, results in a violation of random assignment while simultaneously engaging in “creaming,” making the experiment no more rigorous than a quasi-experiment. Policymakers want to know whether mandating counseling leads to lower rates of recidivism in comparison to other approaches. This question has two parts: (1) Will they attend treatment if mandated? (2) Will treatment have an impact on recidivism if they do attend? Both true and quasi-experiments must grapple with how to tease apart the two parts of this question.

Other limitations include variability across studies concerning what constitutes successful treatment completion. In some cases, the definition was clear (e.g., completing 70% or 80% of the required sessions) and in other studies, it was unspecified. Future researchers should carefully specify what qualifies as successful completion of treatment and also examine the relationship between number of treatment sessions attended and outcome to identify any potential “dose–response” curve. The reliance on dichotomous variables of recidivism may be an overly conservative estimate and dampen the effect size of batterers’ treatment. The overall effect sizes may be larger if one uses a reduction of violence rather than cessation of violence as the outcome measure. However, doing so would result in the inclusion of a smaller number of studies, as several early studies do not report the necessary statistics. In addition, the clinical significance of the change in violence attributable to batterers’ intervention may be questionable.

All longitudinal studies are affected by follow-up rates. As shown in Tables 1 and 2, many studies fail to report participation rates of partners at follow-up. Where partner follow-up contact rates are reported, they range from 22% to 90% of the sample. Those who are lost to follow-up are thought to be more abusive (DeMaris & Jackson, 1987), and therefore success rates may be inflated (Tolman & Bennett, 1990). As such, the resultant effect sizes would also probably be overestimates.

Like partner reports, police reports as outcome measures of recidivism are also problematic and may not adequately reflect reality. With couples already involved in family violence interventions, only about one out of every five domestic violence assaults are reported to the authorities (Rosenfeld, 1992). In some jurisdictions, police reports themselves are inaccurate. Crimes committed outside of the state or local jurisdiction, or incidents of violence in which adjudication was deferred may not appear on the criminal record. Crimes that do appear on the record may be ambiguous as to whether they were family violence or other types of

assault, and researchers have to grapple with which types of crimes “count” in terms of recidivism.

Moreover, the effect size due to treatment for court-mandated batterers is confounded with the strength of the coordinated efforts of the police, probation, and legal system. The potency of the legal system that sanctions men for noncompliance may have a profound effect on treatment completion rates and, as a result, the effect of treatment. Yet, few studies attempt to examine the additive effects of arrest, prosecution, treatment, probation, and legal action for noncompliance (Babcock & Steiner, 1999; Murphy, Musser, & Maton, 1998, are exceptions).

Given these methodological and pragmatic issues, it is not surprising that the effect sizes attributable to batterers’ treatment are small. Although we excluded treatment comparison studies because they only allow an estimate of the size of the difference between two active interventions, the entire literature on batterers’ intervention is actually predominated by component analysis studies, attempting to measure the additive component of the treatment on top of the legal interventions. Since involvement in the legal system is probably beneficial in reducing recidivism (Dutton, 1987), court-ordered treatment programs must reduce abuse recidivism further to demonstrate the effectiveness of treatment over and above legal-system interventions (Rosenfeld, 1992). Differences between two active interventions are more difficult to find than between treatment and no-treatment conditions. Added to that is the spontaneous violence cessation rate in nonclinical samples of about 35% (O’Leary et al., 1989). For batterers’ interventions to be proven effective, they must supercede both the spontaneous recovery rate and the effects of legal interventions.

4.5. Clinical and policy implications

Policymakers should not accept the null hypothesis and dismiss the potential for batterers’ interventions to have an impact on intimate partner abuse. Results showing a small effect of treatment on violence abstinence do not imply that we should abandon our current battering intervention programs. Similar small treatment effects are found in meta-analyses of substance abuse treatments when abstinence from alcohol is the outcome of interest (Agosti, 1995). Yet, some people are able to dramatically transform their lives following substance abuse or battering interventions. Given what we now know about the overall small effect size of batterers’ treatment, the energies of treatment providers, advocates, and researchers alike may best be directed at ways to improve batterers’ treatment. Because no one treatment model or modality has demonstrated superiority over then others, it is premature for states to issue mandates limiting the range of treatment options for batterers. Battering intervention agencies are more likely to improve their services by adding components or tailoring their treatments to specific clientele, than by rigidly adhering to any one curriculum in the absence of empirical evidence of its superior efficacy. Different types of batterers may preferentially benefit from specific forms of interventions (Saunders, 1996), yet no controlled treatment-matching studies have been conducted to date. While a small number of studies have assessed group and couples’ formats, no published studies to date have attempted to assess the efficacy of individual treatment for battering, although this researchers are embarking on this frontier (e.g., Fruzzetti, 2001; Rathus, 2001). Promising directions for improving treatment efficacy

include targeting treatments to specific subsamples, such as different ethnic minority groups, batterers who are chemically dependent, batterers at different motivational stages, different types of batterers (e.g., family-only, borderline, and antisocial/generally violent types), and women arrested for domestic violence. Treatment providers should develop alternative techniques and collaborate with researchers to evaluate their efficacy in an effort to develop evidence-based practice. To this end, researchers need to become an integral part of the coordinated community response to domestic violence.

Batterers' treatment is just one component of the coordinated community response to domestic violence. Police response, prosecution, probation, as well as treatment all affect recidivism of domestically violent partners. Even the best court-mandated treatment programs are likely to be ineffective in the absence of a strong legal response in initial sentencing and in sanctioning offenders who fail to comply with treatment. Even then, treatment may not be the best intervention for all batterers. Alternative sanctions should be developed and empirically tested along with alternative treatments.

References

References marked with an asterisk indicate studies included in the meta-analysis.

- Abel, E. (in press). Treatment for female "batterers." *Journal of Family Violence*.
- Adams, D. A. (1988). Counseling men who batter: A pro-feminist analysis of five treatment models. In M. Bograd, & K. Yllo (Eds.), *Feminist perspectives on wife abuse* (pp. 176–199). Newbury Park, CA: Sage.
- Agosti, V. (1995). The efficacy of treatments in reducing alcohol consumption: A meta-analysis. *International Journal of the Addictions*, 30, 1067–1077.
- Austin, J. B., & Dankwort, J. (1999). Standards for batterer programs: A review and analysis. *Journal of Interpersonal Violence*, 14, 152–168.
- Babcock, J. C., & LaTaillade, J. (2000). Evaluating interventions for men who batter. In J. Vincent, E. Jouriles (Eds.), *Domestic violence: Guidelines for research-informed practice* (pp. 37–77). Philadelphia: Jessica Kingsley.
- *Babcock, J. C., & Steiner, R. (1999). The relationship between treatment, incarceration, and recidivism of battering: A program evaluation of Seattle's coordinated community response to domestic violence. *Journal of Family Psychology*, 13, 46–59.
- Beutler, L. (1991). Have all won and must all have prizes? Revisiting Luborsky et al.'s verdict. *Journal of Consulting and Clinical Psychology*, 59, 226–232.
- Bargarozzi, D. A., & Giddings, C. W. (1983). Conjugal violence: A critical review of current research and clinical practices. *The American Journal of Family Therapy*, 11, 3–15.
- Brannen, S. J., & Rubin, A. (1996). Comparing the effectiveness of gender-specific and couples groups in a court-mandated spouse abuse treatment program. *Research on Social Work Practice*, 6, 405–416.
- Bureau of Justice Statistics (1998). *Violence by intimates: Analysis of data on crimes by current or former spouses, boyfriends, and girlfriends*. Washington, DC: U.S. Department of Justice, NCJ-167237.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin Co.
- Chambless, D. L., Sanderson, W. C., Shoham, V., Johnson, S. B., Pope, K. S., Crits-Cristoph, P., Baker, M., Johnson, B., Woody, S. R., Sue, S., Beutler, L., Williams, D. A., & McCurry, S. (1996). An update on empirically validated therapies. *Clinical Psychologist*, 49, 5–11.
- *Chen, H., Bersani, S., Myers, S. C., & Denton, T. (1989). Evaluating the effectiveness of a court-sponsored abuser treatment program. *Journal of Family Violence*, 4, 309–322.

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Davis, R. C., & Taylor, B. G. (1999). Does batterer treatment reduce violence? A synthesis of the literature. *Women and Criminal Justice*, 10, 69–93.
- *Davis, R. C., Taylor, B. G., & Maxwell, C. D. (1998). Does batterer treatment reduce violence? A randomized experiment in Brooklyn. *Justice Quarterly*, 18, 171–201.
- DeMaris, A., & Jackson, J. K. (1987). Batterers reports of recidivism after counseling. *Social Casework*, 68, 458–465.
- DeRubeis & Crits-Cristoph (1998). Empirically supported individual and group psychological treatment for adult mental disorders. *Journal of Consulting and Clinical Psychology*, 66, 37–52.
- *Dobash, R., Dobash, R. E., Cavanagh, K., & Lewis, R. (1996). Reeducation programs for violent men—An evaluation. *Research Findings*, 46, 309–322.
- Dunford, F. W. (1998, July). *Experimental design and program evaluation*. Paper presentation at the Program Evaluation and Family Violence Research: An International Conference, Durham, NH.
- *Dunford, F. W. (2000). The San Diego Navy experiment: An assessment of interventions for men who assault their wives. *Journal of Consulting and Clinical Psychology*, 68, 468–476.
- Dutton, D. G. (1987). The criminal justice system response to wife assault. *Law and Human Behavior*, 11, 189–206.
- *Dutton, D. G., Bodnarchuk, M., Kropp, R., & Hart, S. D. (1997). Wife assault treatment and criminal recidivism: An 11-year follow-up. *International Journal of Offender Therapy and Comparative Criminology*, 41, 9–23.
- Dutton, D. G. (1998). *The abusive personality: Violence and control in intimate relationships*. New York: Guilford Press.
- Edleson, J. (1996). Controversy and change in batterers' programs. In J. L. Edleson, & Z. C. Eisikovits (Eds.), *Future interventions with battered women and their families* (pp. 154–169). Thousand Oaks, CA: Sage.
- Edleson, J., & Grusznski, R. (1988). Treating men who batter: Four years of outcome data from the domestic abuse project. *Journal of Social Service Research*, 12, 3–22.
- *Edleson, J. L., & Syers, M. (1991). The effects of group treatment for men who batter: An 18-month follow-up study. *Research in Social Work Practice*, 1, 227–243.
- Fagan, J. (1989). Cessation of family violence: Deterrence and dissuasion. In L. Ohlin, & M. Tonry (Eds.), *Family Violence: Crime and Justice: A review of research* (pp. 377–425). Chicago: The University of Chicago.
- *Feder, L., & Forde, D. (1999, July). *A test of the efficacy of court-mandated counseling for convicted misdemeanor domestic violence offenders: Results from the Broward experiment*. Paper presented at the International Family Violence Research Conference, Durham, NH.
- *Flournoy, P.S. (1993). A comparison of groups for men who batter (Doctoral dissertation, Washington State University), *Dissertation Abstracts International*, 53, 5974.
- *Ford, D. A., & Regoli, M. J. (1993). The criminal prosecution of wife batterers: Process, problems, and effects. In N. Z. Hilton (Ed.), *Legal responses to wife assault* (pp. 127–164). Newbury Park, CA: Sage.
- Fruzzetti, A. E. (2001, November). Dialectical Behavior Therapy adapted for treatment of partner-violent men. *Panel discussion at the Association for the Advancement of Behavior Therapy Convention* (Philadelphia, PA.).
- Gelles, R. J. (1980). Violence in the family: A review of research in the seventies. *Journal of Marriage and the Family*, 42, 873–885.
- *Gondolf, E. (1998). Do batterer programs work? A 15-month follow-up of a multi-site evaluation. *Domestic Violence Report*, 3, 64–65, 78–79.
- Gondolf, E. (2000). Reassault at 30-months after batterer programs intake. *International Journal of Offender Therapy and Comparative Criminology*, 44, 111–128.
- Gondolf, E. (2001). Limitation of experimental evaluations of batterer programs. *Trauma, Violence, and Abuse*, 2, 79–88.
- Gondolf, E. W. (1997). Patterns of reassault in batterer programs. *Violence and Victims*, 12, 373–387.
- Guerney, B. G. (1977). *Relationship enhancement: Skill training programs for therapy, problem prevention, and enrichment*. San Francisco: Jossey-Bass.

- *Hamberger, K., & Hastings, J. (1988). Skills training for treatment of spouse abusers: An outcome study. *Journal of Family Violence*, 3, 121–130.
- Hamberger, K., & Hastings, J. (1993). Court-mandated treatment of men who assault their partner: Issues, controversies, and outcomes. In N. Z. Hilton (Ed.), *Legal responses to wife assault* (pp. 188–229). Newbury Park, CA: Sage.
- *Harrell, A. (1991). *Evaluation of court ordered treatment for domestic violence offenders* (Final Report). Washington, DC: Urban Institute.
- *Hawkins, R., & Beauvais, C. (1985, August). *Evaluation of group therapy with abusive men: The police record*. Paper presented at the meeting of the American psychological association, Los Angeles, CA.
- Healey, K., Smith, C., O'Sullivan, C. (1998). *Batterer intervention: Program approaches and criminal justice strategies*. Report to the National Institute of Justice, Washington, DC.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Heinsman, D. T., & Shadish, W. R. (1996). Assignment methods in experimentation: When do nonrandomized experiments approximate answers from randomized experiments? *Psychological Methods*, 1, 154–169.
- Holtzworth-Munroe, A., Beatty, S. B., & Anglin, K. (1995). The assessment and treatment of marital violence: An introduction for the marital therapist. In N. S. Jacobson, & A. S. Gurman (Eds.), *Clinical handbook of couple therapy* (pp. 317–339). New York: Guilford Press.
- Huffcutt, A. I., & Arthur, W. (1995). Development of a new outlier statistic for meta-analytic data. *Journal of Applied Psychology*, 80, 327–334.
- Jacobson, N. S. (1993, October). *Domestic violence: What are the marriages like?* Plenary delivered at the annual meeting of American Association of Marriage and Family Therapy, Anaheim, CA.
- Johnson, B. T. (1995). *D-Stat: Software for the meta-analytic review of research literatures* (Version 1.11). Hillsdale, NJ: Erlbaum.
- *Leong, D. J., Coates, C. J., & Hoskins, J. (1987). *Follow-up of batterers treated in a court-ordered treatment program*. Paper presented at the Third National Family Violence Research Conference, University of New Hampshire, Durham, NH.
- Levesque, D. A., & Gelles, R. J. (1998, July). *Does treatment reduce recidivism in men who batter? A meta-analytic evaluation of treatment outcome*. Paper presentation at the Program Evaluation and Family Violence Research: An International Conference, Durham, NH.
- Lipchik, E., Sirles, E. A., & Kubicki, A. D. (1997). Multifaceted approaches in spouse abuse treatment. In R. Geffner, S. B. Sorenson, & P. K. Lundberg-Love (Eds.), *Violence and sexual abuse at home: Current issues in spousal battering and child maltreatment* (pp. 131–148). New York: Haworth Press.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist*, 48, 1181–1209.
- Loesel, F., & Koeferl, P. (1987). Evaluationsforschung zur sozialtherapeutischen Anstalt: Eine meta-analyse (Evaluation research on the social-therapeutic prison: A meta-analysis). *Gruppendynamik*, 18, 385–406.
- Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies: Is it true that “everyone has won and all must have prizes”? *Archives of General Psychiatry*, 32, 995–1008.
- McCartney, K., & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. *Child Development*, 71, 173–180.
- Miller, W. R., & Rollnick, S. (1991). *Motivational interviewing: Preparing people to change addictive behavior*. New York: Guilford Press.
- Moffit, T. E., Caspi, A., Krueger, R. F., Magdol, L., Margolin, G., Silva, P. A., & Sydney, R. (1997). Do partners agree about abuse in their relationship? A psychometric evaluation of interpartner agreement. *Psychological Assessment*, 9, 47–56.
- *Morrel, T. M., Elliott, J. D., Murphy, C. M., & Taft, C. (2003). A comparison of cognitive-behavioral and supportive group therapies for male perpetrators of domestic abuse. *Behavior Therapy*, 24, 77–95.
- Murphy, C. M., & Baxter, V. A. (1997). Motivating batterers to change in the treatment context. *Journal of Interpersonal Violence*, 12, 607–619.

- *Murphy, C. M., Musser, P. H., & Maton, K. I. (1998). Coordinated community intervention for domestic abusers: Intervention system involvement and criminal recidivism. *Journal of Family Violence*, 13, 263–284.
- *Newell, R. G. (1994). The effectiveness of court-mandated counseling for domestic violence: An outcome study. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 53, 1193.
- O'Leary, K. D., Barling, J., Aria, I., Rosenbaum, A., Malone, J., & Tyree, A. (1989). Prevalence and stability of physical aggression between spouses. A longitudinal analysis. *Journal of Consulting and Clinical Psychology*, 57, 263–268.
- O'Leary, K. D., Heyman, R. E., & Neidig, P. H. (1999). Treatment of wife abuse: A comparison of gender-specific and conjoint approaches. *Behavior Therapy*, 30, 475–506.
- *Palmer, S. E., Brown, R. A., & Barrera, M. E. (1992). Group treatment program for abusive husbands: Long term evaluation. *American Journal of Orthopsychiatry*, 62, 276–283.
- Peluch, J. (1987). *Goals for men's treatment groups*. Unpublished manuscript, Family Services of Hamilton-Wentworth, Ontario, Canada.
- Pence, E., & Paymar, M. (1993). *Education groups for men who batter: The Duluth model*. New York: Springer-Verlag.
- Rathus, J. H. (2001, November). *Dialectic behavior therapy adapted for the treatment of partner violent men*. Panel presentation at the Association for the Advancement of Behavior Therapy, Philadelphia, PA.
- Rosenberg, M. S., Adams, D. C., & Gurevitch, J. (1997). Metawin: Statistical software for meta-analysis with resampling tests. Sunderland, MA: Sinauer Associates.
- Rosenfeld, B. D. (1992). Court ordered treatment of spouse abuse. *Clinical Psychology Review*, 12, 205–226.
- Rosenthal, R. (1995). Writing meta analytic reviews. *Psychological Bulletin*, 118, 183–192.
- Rosnow, R. L., & Rosenthal, R. (1988). Focused tests of significance and effect size estimation in counseling psychology. *Journal of Counseling Psychology*, 35, 203–208.
- Saunders, D. G. (1996). Feminist–cognitive–behavioral and process–psychodynamic treatments for men who batter: Interaction of abuser traits and treatment model. *Violence and Victims*, 11, 393–414.
- Shadish, W. R., & Ragsdale, K. (1996). Random versus nonrandom assignment in controlled experiments: Do you get the same answer? *Journal of Consulting and Clinical Psychology*, 64, 1290–1305.
- Sherman, L. W., Smith, D., Schmidt, J. D., & Rogan, D. P. (1992). Crime, punishment, and stake in conformity: Legal and informal control of domestic violence. *American Sociological Review*, 57, 680–690.
- Smith, M. L., Glass, G. V., & Miller, T. (1980). *The benefits of psychotherapy*. Baltimore: Johns Hopkins University Press.
- Sonkin, D. J., Martin, D., & Walker, L. E. A. (1985). *The male batterer: A treatment approach*. New York: Springer-Verlag.
- *Stacey, W. A., & Shupe, A. (1984). *An evaluation of three programs for abusive men in Texas* (Research Monograph No. 29). Arlington, TX: Center for Social Research, University of Texas.
- Straus, M. A. (1979). Measuring intrafamily conflict and violence: The Conflict Tactics (CT) Scales. *Journal of Marriage and the Family*, 41, 75–88.
- Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The revised Conflict Tactics Scales (CTS-2). *Journal of Family Issues*, 17, 283–316.
- Sullivan, C. M., & Bybee, D. I. (1999). Reducing violence using community based advocacy for women with abusive partners. *Journal of Consulting and Clinical Psychology*, 67, 43–53.
- *Taft, C. T., Murphy, C. M., Elliot, J. D., & Morrel, T. M. (1990). Attendance enhancing procedures in group counseling for domestic abusers. *Journal of Counseling Psychology*.
- Tolman, R. T., & Bennett, L. (1990). A review of quantitative research on men who batter. *Journal of Interpersonal Violence*, 5, 87–118.
- U.S. Attorney General (1994). *Task Force on Family Violence: Final report*. U.S. Department of Justice.
- Waldo, M. (1985). A curative factor framework for conceptualizing group counseling. *Journal of Counseling and Development*, 64, 52–58.
- *Waldo, M. (1988). Relationship enhancement counseling groups for wife abusers. *Journal of Mental Health Counseling*, 10, 37–45.

- Wang, M. C., & Bushman, B. J. (1998). Using the normal quantile plot to explore meta-analytic data sets. *Psychological Methods*, 3, 46–54.
- Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A., & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin*, 117, 450–468.
- Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Beverly Hills, CA: Sage.
- Zimring, F. E. (1989). Toward a jurisprudence of family violence. In L. Ohlin, & M. Tonry (Eds.), *Family violence: Crime and justice: A review of research* (pp. 547–569). Chicago: University of Chicago Press.

Research on Social Work Practice

<http://rsw.sagepub.com/>

Comparing the Effectiveness of Gender-Specific and Couples Groups in a Court-Mandated Spouse Abuse Treatment Program

Stephen J. Brannen and Allen Rubin

Research on Social Work Practice 1996 6: 405

DOI: 10.1177/104973159600600401

The online version of this article can be found at:

<http://rsw.sagepub.com/content/6/4/405>

Published by:



<http://www.sagepublications.com>

Additional services and information for *Research on Social Work Practice* can be found at:

Email Alerts: <http://rsw.sagepub.com/cgi/alerts>

Subscriptions: <http://rsw.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://rsw.sagepub.com/content/6/4/405.refs.html>

Comparing the Effectiveness of Gender-Specific and Couples Groups in a Court-Mandated Spouse Abuse Treatment Program

Stephen J. Brannen

Uniformed Services University of the Health Sciences

Allen Rubin

University of Texas at Austin

This study represents the first attempt to directly compare two common methods of providing spouse abuse intervention, group treatment of couples, or gender-specific groups. Forty-nine couples were randomly assigned to one of the two treatments. Of these, data were available at posttest on only 42 couples because of attrition of the victimized spouses. A multiple analysis of covariance at posttest demonstrated that for the majority of abusers it did not matter which treatment was used. However, for those with a history of alcohol abuse, the couples approach was clearly superior. Analysis of victims' reports at a 6-month follow-up suggests that neither treatment approach was more effective in sustaining initial treatment gains over time. Finally, the issue of victim safety was addressed. Qualitative assessment of weekly reporting sheets suggests that women who received the couples group intervention were in no more danger than those receiving treatment in the gender-specific groups.

Spouse abuse is a problem of growing concern within the United States. Straus and Gelles (1986, 1988, 1990) estimated that approximately 1.8 million women annually are physically abused by their husbands and that perhaps as many as 60% of all wives have been victimized at some time during their marriage. In fact, it is estimated that in the United States, one woman is physically abused every 12 seconds (Stark, Flitcraft, & Frazier, 1979), raped every 46 seconds (Kilpatrick, 1992), and murdered by a partner every 6 hours (American College of Obstetricians and Gynecologists, 1989). Programs aimed at treating this serious problem have been in existence since the mid-1970s, but more controlled studies are needed to evaluate the

Authors' Note: Correspondence may be addressed to Stephen J. Brannen, Department of Family Medicine, Uniformed Services University of the Health Sciences, 4301 Jones Bridge Road, Bethesda, MD 20814-4789.

Research on Social Work Practice, Vol. 6 No. 4, October 1996 405-424
© 1996 Sage Publications, Inc.

effectiveness of the interventions they employ (Caesar & Hamberger, 1989; Edleson & Tolman, 1992; Sonkin, Martin, & Walker, 1985).

Interventions primarily have been based on two dominant perspectives as to the origins, maintenance, and treatment of domestic violence. One perspective argues for the use of gender-specific groups aimed at identifying and eliminating abusive behavior within the batterer. The other—the family systems/social learning theory perspective—argues for the treatment of the abusive relationship using couples group therapy. Both perspectives utilize cognitive-behavioral intervention strategies as the basis of treatment. Numerous quasiexperimental studies have been conducted to evaluate the effectiveness of both the gender-specific group and the couples group approaches; however, methodological problems have limited the generalizability of the results (Edleson & Tolman, 1992). To date, no study has been conducted directly comparing these two dominant intervention approaches. The purpose of this article is to present the findings of a study that directly compares these two interventions.

HYPOTHESIS

The existing literature provides no compelling rationale for predicting whether either treatment approach is more effective than the other with spouse abusers in general. This study, however, was limited to intact couples (couples who resided together and who indicated a desire to remain in their current relationship). The vast majority of cases in the study (66.7%) involved relatively minor incidents of abuse (i.e., pushing, shoving, grabbing, slapping, holding a partner down), whereas 33.3% involved severe physical abuse (i.e., punching, kicking, choking, use of a weapon). For this population, the couples group intervention appeared to offer better prospects of effectiveness than the gender-specific group intervention.

The marital therapy literature suggests that communication issues are treated most effectively in conjoint therapy (Gurman & Kniskern, 1978; Gurman, Kniskern, & Pinsof, 1986; Margolin, 1979; Ohlsen, 1979). Spouse abuse occurs in the dyadic context, and interactional factors are involved in the escalation of conflict (Deschner, 1984; Geller, 1982; Neidig, 1984; Neidig & Friedman, 1984). If couples wish to remain together, then they must learn communication and problem-solving skills. Without these skills, future conflicts may escalate into a relapse of abusive behavior (Margolin, 1979; Neidig, 1984). Family systems and social learning theories suggest that even if physical aggression terminates as a result of treatment, unless other

relationship factors such as psychological abuse and general marital discord are also addressed, the couple may be at risk for long-term relapse of physical violence, psychological aggression, or intimidation. Moreover, because many victimized women return to their spouses, the couples group intervention offers the dyad an opportunity to work on developing improved patterns of communication and conflict management (Ohlsen, 1979).

In light of this reasoning the hypothesis for this study was that the couples group intervention would be more effective than the gender-specific group intervention in achieving the following five outcome objectives:

1. Enhancing the clients' ability to positively resolve conflicts within the dyadic relationship;
2. Reducing the level of violence within the dyadic relationship;
3. Enhancing the level of communication within the dyadic relationship;
4. Enhancing the level of marital satisfaction within the relationship; and
5. Preventing recidivism.

This study also examined the relative effectiveness of the two approaches when controlling for alcohol abuse. Alcohol abuse has been identified as a key factor involved in domestic violence situations. In fact, experts suggest that alcohol use or alcohol abuse is present in as many as 50% of all cases of domestic violence (Edleson & Tolman, 1992; Flanzer, 1993; Gelles & Cornell, 1990; Neidig, 1984; Ptacek, 1988). For purposes of this study, a history of alcohol abuse is defined as a positive identification by the subject on the screening form of alcohol as a recurrent problem in the marriage.

METHOD

Clients and Group Assignment

All clients were intact couples who had indicated a desire to remain in their current relationship and who were referred by the 9th County Court at Law, San Antonio, Bexar County, Texas. Pretesting was administered after screening and assessment interviews were completed. The clients were then randomly assigned to one of the two treatment conditions: the couples group intervention or the gender-specific group intervention. Sixty couples met all inclusionary criteria at the time of the intake interview and were referred for screening. Of these 60 couples, 11 declined to participate and were referred to other treatment programs approved by the court. This left a total of 49 couples who participated in the study.

Although 49 probationers completed the program, pretest to posttest comparisons are based on the data for only 42 individuals because 7 spouses dropped out during the intervention phase. Of the 7, 6 were women assigned to the gender-specific group intervention, whereas 1 was assigned to the couples group intervention. The sample characteristics of those probationers who completed the program are displayed in Table 1. Preliminary analyses indicated that there were no significant differences between groups on key demographic variables (age, income, education, employment status, ethnicity, marital status). There was a significant difference between the groups on the years they had been together. The mean years together was 7.88 for the couples group and 5.39 for the gender-specific group. However, when years together was entered into the analysis as a covariate, the difference between the two interventions was not significant, Wilks's $\Lambda = .769$, $F(6, 22) = 1.1$, $p = .391$.

Interventions

Couples Group Intervention

The couples group intervention assessed in this study is based on a model developed by Neidig and Friedman (1984). The intervention uses a cognitive-behavioral approach with a core curriculum designed to enable clients to accept personal responsibility for violent behavior, contract for a commitment to change, develop and use time-out and other security mechanisms, understand the unique factors involved in the violence sequence, master anger control techniques, and develop the ability to contain interpersonal conflict through the problem-solving process (p. 9). In addition to the core curriculum, specific anger control techniques were taught, which include assertion training, stress-inoculation training, empathy building, coping with criticism, stress-management training, communication skills, sex-role stereotyping, marital dependency, and isolation and social support. The approach included three basic components during each session: instruction, behavioral rehearsal, and feedback. The focus remained on establishing a strong relationship and eliminating violence within the relationship rather than on making any wide-sweeping changes in the individuals' personalities.

Gender-Specific Group Intervention

The gender-specific group intervention used in this study was based on a model developed at the Domestic Abuse Project (DAP), Minneapolis, Min-

TABLE 1: Demographic Characteristics of the Sample

<i>Characteristic</i>	<i>Couples Group n = 22</i>	<i>Gender-Specific Group n = 26</i>
Age		
Mean	32.18	30.85
SD	6.52	6.82
Range	19-46	19-42
Education		
Mean	12.43	11.69
SD	1.72	2.09
Range	8-17	10-15
Employment status		
Employed	17 (77.3)	19 (73.1)
Unemployed	5 (22.7)	7 (26.9)
Ethnicity		
African American	2 (9.1)	2 (7.4)
Anglo	8 (36.4)	3 (11.1)
Asian	1 (4.5)	1 (3.7)
Hispanic	11 (50)	21 (77.8)
Total family income		
Mean	\$19,701	\$19,843
SD	\$13,847	\$16,980
Range	\$3,000-\$58,000	\$6,500-\$85,000
Marital status		
Married	16 (76.2)	15 (71.4)
Cohabiting	5 (23.8)	6 (28.6)
Length of current relationship (in years)		
Mean	7.88	5.39*
SD	5.48	3.93
Range	1-19	1-17
Religion		
Catholic	14 (66.7)	16 (61.5)
Muslim	1 (4.8)	1 (3.8)
Protestant	6 (28.6)	6 (23.1)
None	1 (4.8)	3 (11.5)
Alcohol abuse a problem ^a		
Yes	5 (22.7)	7 (25.9)
No	17 (77.3)	20 (74.1)

NOTE: Numbers in parentheses are percentages.

a. Alcohol is listed here because it was used as a control variable in the factorial design.

* $p < .05$.

nesota, as described in their *Men's Treatment Handbook* (Rusinoff, 1990). The DAP model emphasizes the male as perpetrator and responsible party for the abuse within the marriage. The primary focus of programs of this type is on modifying the abusive behavior of batterers. According to the DAP,

work with victims is supplemental to that of the batterers. The DAP men's program is a 10-week program, which is separated into educational and process sessions. This study adapted the DAP intervention to consist of 12 weekly sessions, each session lasting 1-1/2 hours. This modification was made to make both interventions used in the study consistent. Both educational and process materials were incorporated into weekly sessions. The DAP's cognitive-behavioral approach emphasizes instruction in the definition of violence, origins of aggression perpetrated against women, and power issues. The context in which treatment takes place is increasing the safety of the victim. It focuses on establishing accountability in the men who batter.

Group treatment for victims is viewed as "supplemental to, or complementary to the perpetrators' groups" (Ganley, 1989). The purpose of treatment with the victims is to develop within them a sense of empowerment and to further develop the victims' ability to protect themselves. The program attempts to eliminate complete reliance on the marital partner by assisting the women in developing new positive social support systems, which provide for the free expression of emotion and that further break down the social isolation often experienced by these women (Ganley, 1989). In this study, the women's group, based on the cognitive-behavioral model developed by the DAP (DAP Women's Therapy Team, 1991), included sessions dealing with the myths and beliefs associated with violence, progression of violence, sociocultural factors leading to spouse battering, power and control issues within a relationship, victims' survival skills, protection planning, anger management techniques, stress management, and communications techniques.

Outcome Measures

The first two outcome objectives—dealing with conflict resolution ability and level of violence—were measured using the three subscales of the Modified Conflict Tactics Scale (MCTS) (Neidig, 1986). The MCTS, a self-report instrument, assesses the frequency of various conflict resolution tactics in the relationship. Spouses rate their own and their partner's behavior on each of 24 items that refer to possible areas of family conflict. The MCTS yields three subscales that address physical and verbal aggression tactics used by couples in resolving conflict. In addition, the MCTS contains a reasoning subscale consisting of three questions that address the individual's attempts at resolving conflicts in a positive fashion. Each item is rated on a scale from 1 (*never*) to 7 (*more than 20 times*), based on how often the event has occurred over a specified time period (from pretest to posttest). The MCTS subscales have demonstrated high internal consistency reliability. Alpha coefficients

based on women scores reporting on their male partners are high for all scales: physical abuse, .92; severe physical abuse, .86; and psychological abuse, .84. Alpha coefficients are not provided for the reasoning scale because it consists of only three questions (Pan, Neidig, & O'Leary, 1994). Estimates of internal consistency (Cronbach's alpha) in this study for the females reporting on their mates are .84 (psychological abuse), .80 (physical abuse), .64 (severe physical abuse), and .58 (reasoning).

The third outcome objective—level of communication within the dyadic relationship—was assessed using the communication subscale of the McMaster Family Assessment Device (FAD, Version 3) (Epstein, Baldwin, & Bishop, 1983). The FAD-III is a self-report instrument that assesses family functioning on seven dimensions, including communications. The communication subscale is comprised of nine items and has demonstrated adequate internal consistency and test-retest reliability (.75 and .72, respectively) (Miller, Epstein, Bishop, & Keitner, 1985). In this study, a reliability analysis (Cronbach's alpha) of the communication subscale yielded a coefficient of .58.

The fourth outcome objective—level of marital satisfaction—was measured using the Marital Satisfaction Inventory (MSI) (Snyder, 1981). The MSI is a multidimensional self-report measure that identifies for each spouse the nature and extent of marital distress along several key dimensions of their relationship. For this study, only the global distress scale (GDS) was used. The GDS contains items measuring the individual's overall dissatisfaction with the marriage. The instrument has been shown to have high internal consistency (.97) and test-retest reliability (.92) (Snyder, 1981). A reliability analysis (Cronbach's alpha) of the GDS using this study yielded a coefficient of .94.

The fifth outcome objective—preventing recidivism—was measured using the Long-Term Evaluation Form (adapted from Neidig & Friedman, 1984) and confirmed by a review of police and probation officer records to determine if any of the subjects had come into further contact with either the police or court. The Long-Term Evaluation Form was administered over the telephone at a 6-month posttreatment follow-up and included questions concerning any further problems with violence after the cessation of treatment; the number, seriousness, and nature of further incidents; and the use of any law enforcement intervention.

Safety Issues

In response to concerns levied against the use of couples group interventions as posing excessive risks to the safety of the abuse victim (Edleson &

Tolman, 1992; Harris, 1986, Rosenbaum & O'Leary, 1986), an elaborate safety net was established to ensure that none of the women were placed into a position of receiving further physical or psychological abuse as a result of their participation in this study. During the planning phase of the study, each of the program directors for the several existing court-affiliated programs for batterers was advised of the scope of the current study. These directors identified potential threats to the safety of the women involved and suggested mechanisms to reduce the risk of further violence directed at the victims as a result of their involvement in the program. In addition, the clinical services director of the women's shelter was advised of the study and provided guidance on the shelter's emergency referral procedures.

A separate orientation was provided for the victims. During this orientation, the names and phone numbers of the respective group facilitator and the primary investigator were provided, including the investigator's 24-hour emergency phone number. Participants were encouraged to use the phone numbers at any time they felt threatened in the relationship. The victims were also provided the phone numbers of local law enforcement officials, the numbers to the battered women's shelter, and instructions on how to make an emergency access to the shelter.

Data were gathered to monitor safety using the Treatment Project Weekly Summary (TPWS) (Neidig, 1992). This questionnaire, a modified version of the MCTS, was administered weekly to both spouses. The questionnaire contains items concerning the continued use of psychological and physical abuse and a series of brief questions concerning any disagreements that led to any violence, and whether these issues were raised during treatment sessions.

These summaries were returned to the group facilitators on a weekly basis, were reviewed at the beginning of each session, and were forwarded to the principal investigator. If any indication became known during the course of a session, the facilitator immediately telephoned the principal investigator to advise him of the nature of the threat situation. A follow-up phone call was made to the victim to establish the existence of a threat, and to provide additional information on how to receive emergency interventions as needed. In addition, the principal investigator completed an audit of all TPWSs.

Facilitators

Group facilitators were certified social workers recruited from the local community. Each selected facilitator was assigned to the intervention method for which he or she had expressed an interest in working. To ensure that

facilitators adhered to the treatment protocols, all sessions were audiotape-recorded, with 20% of the tapes randomly sampled on a weekly basis.

RESULTS

Pretest-Posttest Phase

A 2×2 multivariate analysis of covariance was performed on six dependent variables: level of reasoning, level of psychological abuse, level of physical abuse, level of severe physical abuse, level of communications, and level of marital satisfaction.

Although data were collected from both the victims and probationer regarding the levels of reasoning, psychological abuse, physical abuse, and severe physical abuse, only victims' data were used in the present study because previous research indicated that women more often report violence than do their male partners (Edleson & Brygger, 1986; Edleson & Syers, 1990; Jouriles & O'Leary, 1985; Szinovacz, 1983).

SPSS MANOVA (SPSS for Windows, Version 6) was used for the analysis with unique adjustment for nonorthogonality. Order of entry of independent variables was treatment condition, then history of alcohol abuse. Six outliers were adjusted using procedures outlined by Stevens (1992) and Tabachnick and Fidell (1989). Scores on the MCTS were highly skewed. This skewness is to be expected because the sample represents individuals who had been placed on probation for spouse abuse by the county court. Nonetheless, to proceed with MANOVA, transformations were undertaken to change the shape of the distributions of the variables to more nearly normal.

The natural logarithm (LN) method—appropriate for substantially positively skewed data (Stevens, 1992; Tabachnick & Fidell, 1989)—was employed in transforming the scores on the reasoning and psychological abuse subscales of the MCTS. The Base 10 logarithm (LG10) method—appropriate for seriously positively skewed data (Stevens, 1992; Tabachnick & Fidell, 1989)—was used in transforming the scores on the physical aggression and severe physical abuse subscales of the MCTS. When the transformation is applied to distributions in which there are values less than 1, a constant is added to each score so that the smallest value becomes 1 (Tabachnick & Fidell, 1989). Considering that several of the scores on the subscales of the MCTS resulted in scores less than 1, a constant of 1 was added to the scores. The addition of this constant had no negative effects on the data. Once transformations were made, the results of assumptions of normality, homo-

geneity of variance-covariance matrices, linearity, and multicollinearity were satisfactory. There were no significant differences between groups on any of the dependent variables at pretest.

With the use of the Wilks's lambda criterion, there was a significant overall main effect for the treatment factor, Wilks's $\Lambda = .614$, $F(6, 26) = 2.72$, $p = .035$. A moderate association was found between the combined dependent variables and the main effect of treatment, $\eta^2_p = .39$. In addition, the treatment by alcohol use interaction effect was also significant, Wilks's $\Lambda = .521$, $F(6, 26) = 3.99$, $p = .006$. The association between the combined dependent variables and the treatment by alcohol use interaction effect was larger yet, $\eta^2_p = .48$. The main effect for alcohol use, Wilks's $\Lambda = .789$, $F(6, 26) = 1.16$, $p = .356$, was not significant.

When, in a factorial design, both a main effect and interaction effect are significant, the univariate F tests must be reviewed to identify which of the dependent variables are significant. If the same dependent variables are those that account for the significance of both the main effect and the interaction effect, only the interaction effect is interpreted (Stevens, 1992; Tabachnick & Fidell, 1989). In the present study, the significant univariate F s for physical abuse and severe physical abuse accounted for the significant multivariate findings in both the main treatment effect and the interaction effect. Consequently, only the univariate F s associated with the interaction effect were interpreted. An F table for all covariates, main effects, and interaction effects is provided (see Table 2).

The potential problem of an inflated Type I error rate was corrected for by using a Bonferroni-type adjustment to set the critical level of significance for each univariate comparison. The conventional alpha level of .05 was therefore divided by the total number of dependent variables. This adjusted the alpha level at $.05/6 = .008$. All analyses were conducted at this more stringent alpha level.

The following section reports each univariate analysis in detail. Group means and standard deviations at pretest and posttest for all outcome variables are presented in Table 3.

The MCTS Reasoning Scale

A nonsignificant F ratio was found on the reasoning subscale of the MCTS, $F(1, 31) = 1.16$, $p = .289$. Thus no significant differences were found between those subjects assigned to either the couples group or the gender-specific group, as reported by the victims, when controlling for a history of alcohol abuse.

TABLE 2: Test of Covariates, Treatment, History of Alcohol Abuse, and Interaction Effect

<i>Effect</i>	<i>Dependent Variables</i>	<i>Univariate F</i>	<i>df</i>	<i>p</i>
Covariates	Reasoning	1.93	7/31	.099
	Psychological abuse	3.98	7/31	.003**
	Physical abuse	1.85	7/31	.113
	Severe physical abuse	1.07	7/31	.403
	Communication (FAD)	2.36	7/31	.047
	Marital satisfaction	4.13	7/31	.003**
Treatment	Reasoning	0.83	1/31	.370
	Psychological abuse	2.50	1/31	.124
	Physical abuse	8.51	1/31	.007*
	Severe physical abuse	5.00	1/31	.033
	Communication (FAD)	0.89	1/31	.768
	Marital satisfaction	4.79	1/31	.036***
Alcohol abuse	Reasoning	0.41	1/31	.525
	Psychological abuse	4.15	1/31	.050
	Physical abuse	3.92	1/31	.057
	Severe physical abuse	2.37	1/31	.134
	Communication (FAD)	2.13	1/31	.155
	Marital satisfaction	0.22	1/31	.639
Treatment by alcohol abuse interaction	Reasoning	1.16	1/31	.289
	Psychological abuse	4.30	1/31	.047
	Physical abuse	16.84	1/31	.000**
	Severe physical abuse	17.59	1/31	.000**
	Communication (FAD)	0.07	1/31	.794
	Marital satisfaction	1.04	1/31	.315

*Would be interpreted except the interaction effect is also significant.

** $p < .008$.

***Contributes to a significant main effect for treatment but not significant at .008 level.

The MCTS Psychological Abuse Scale

No significant differences were found between groups on the MCTS psychological abuse subscale, as reported by the victims, $F(1, 31) = 4.30$, $p = .047$. Thus, after having controlled for a history of alcohol abuse, the intervention to which the subjects were assigned did not appear to make a significant difference in reducing the level of psychological abuse used as a conflict tactic within the marital relationship.

The MCTS Physical Abuse Scale

There was a significant difference between groups on the scores of the MCTS physical abuse subscale, as reported by the victims, $F(1, 31) = 16.84$,

TABLE 3: Mean Scores and Standard Deviations of Pretest and Posttest Measures of Dependent Variables by Treatment Condition Controlling for History of Alcohol Abuse

	<i>CTS1^a</i>	<i>CTS2^b</i>	<i>CTS3^c</i>	<i>CTS4^d</i>	<i>FAD-III^e</i>	<i>MSIGDS^f</i>
Pretest scores						
Couples group						
No alcohol history						
Mean	1.84	3.51	0.86	0.50	21.31	20.06
SD	1.21	1.24	0.70	0.54	2.96	14.26
Alcohol history						
Mean	1.55	4.50	1.33	0.83	20.60	22.80
SD	1.61	0.68	0.39	0.55	3.91	12.05
Gender-specific group						
No alcohol history						
Mean	1.80	3.33	1.08	0.68	22.78	21.88
SD	1.04	1.18	0.52	0.39	4.27	9.44
Alcohol history						
Mean	1.72	4.49	1.80	1.39	20.33	17.83
SD	0.70	0.55	0.19	0.14	2.94	11.13
Posttest scores^g						
Couples group						
No alcohol history						
Mean	2.26	2.50	0.45	0.27	20.26	19.00
SD	0.58	1.11	0.40	0.38	2.06	9.22
Alcohol history						
Mean	2.34	2.58	0.21	0.00	18.11	14.13
SD	0.74	1.31	0.29	0.00	3.36	11.97
Gender-specific group						
No alcohol history						
Mean	2.28	2.39	0.34	0.07	20.32	23.41
SD	0.83	1.20	0.42	0.17	3.77	10.77
Alcohol history						
Mean	1.86	3.87	1.18	0.76	18.78	24.93
SD	0.37	0.71	0.25	0.47	4.79	12.01

a. Represents transformed MCTS reasoning subscale scores.

b. Represents transformed MCTS psychological abuse subscale scores.

c. Represents transformed MCTS physical abuse subscale scores.

d. Represents transformed MCTS severe physical violence subscale scores.

e. Represents Family Assessment Device, communication scale scores.

f. Represents Marital Satisfaction Inventory, global distress scale scores.

g. Posttest scores are adjusted for effects of covariates.

$p = .000$. Low scores on this MCTS represent lower levels of conflict tactics. Those assigned to the couples group, regardless of any history of alcohol abuse, scored significantly lower (i.e., better) on this measure than those with

a history of alcohol abuse assigned to the gender-specific intervention. However, for those with no history of alcohol abuse, there were no significant difference between interventions. Those assigned to the gender-specific group with no history of alcohol abuse also scored significantly lower than those assigned to the same treatment condition with a history of alcohol abuse. Approximately 35% of the variance in the scores on the MCTS physical abuse subscale among the groups at posttest can be explained by the treatment after controlling for a history of alcohol abuse, using the η^2_p statistic.

The MCTS Severe Physical Abuse Scale

There was a significant difference between the groups on scores of the MCTS severe physical abuse subscale, as reported by the victims, $F(1, 31) = 17.58, p = .000$. Again, those assigned to the couples group, regardless of any history of alcohol abuse, scored significantly lower (i.e., better) on this measure than those with a history of alcohol abuse assigned to the gender-specific intervention. Those assigned to the gender-specific group with no history of alcohol abuse also scored significantly lower than those assigned to the same treatment condition with a positive history of alcohol abuse, as well as scoring lower than either of the two levels of the couples group. Using the η^2_p statistic, approximately 36% of the variance in the scores on the MCTS severe physical abuse scale among the groups at posttest can be explained by the treatment after controlling for a history of alcohol abuse.

Communication

At posttest, a nonsignificant F ratio was found on the communication subscale of the FAD-III (Epstein et al., 1983), $F(1, 31) = 0.07, p = .794$. Thus no significant differences were found between those subjects assigned to either the couples group or the gender-specific group when controlling for a history of alcohol abuse on this measure.

Marital Satisfaction

At posttest, a nonsignificant F ratio, $F(1, 31) = 1.04, p = .315$, was found on the MSI Global Distress Scale (Snyder, 1981). Thus no significant differences were found between those subjects assigned to either the couples group or the gender-specific group when controlling for a history of alcohol abuse on this measure.

Differences Between Ethnic Groups

As has been discussed, Hispanics are overly represented in this study. To verify that ethnicity did not overly influence the results of the omnibus multivariate results, the posttest scores on all measures were examined for differences that might be attributable to ethnicity. Using a multivariate analysis of covariance, no significant main effect was found on the posttest scores of the dependent variables, Wilks's $\Lambda = .4615$, $F(18, 79.68) = 1.3918$, $p = .159$. Subsequent reviews indicated that there were no significant univariate F s for any of the dependent variables.

Follow-Up Phase

Approximately 6 months posttreatment, 26 of the 42 couples who completed the program were located and interviewed. This response rate, although seemingly low, represents 62% of those who completed the program. This response rate is comparable to that achieved in previous studies involving follow-up of spouse abuse studies (Edleson & Syers, 1990; Sherman & Berk, 1984). These 26 included 12 who had been assigned to the couples group and 14 who had been assigned to the gender-specific group. Telephone interviews were conducted by the principal investigator. The primary purpose of this interview was to establish whether any further episodes of physical violence had occurred in the 6 months subsequent to the termination of treatment. As reported by the victims, the probationers were either coded as violent or nonviolent. The use of categorical data required the use of the χ^2 statistic. A two-way cross-tabulation was executed to examine the relationship between the dependent variable (recidivism) and the independent variable (intervention method). This relationship is presented in Table 4.

Table 4 indicates that at 6-month follow-up, the rate of recidivism was independent of the treatment condition to which each of the subjects was assigned. Eleven (91.7%) of the subjects who attended the couples group were reported to have not been physically violent within the 6 months posttreatment, whereas only 1 (8.3%) was reported to have been violent. In the gender-specific group, 13 (92.9) of the subjects remained violence-free in the past 6 months, whereas again only 1 (7.1%) subject was reported to have used violence in the relationship.

To confirm the above results, probation officers provided a review of the probationers' files to identify which, if any, of the subjects' files contained evidence of recidivism, necessitating further interventions on the part of police or the judicial system. The probation officer reports identified one

TABLE 4: Victims' Reports of Probationer Use of Violence at 6-Month Follow-up

<i>Treatment</i>	<i>Report</i>			
	<i>Nonviolent</i>		<i>Violent</i>	
	<i>n</i>	<i>Percentage</i>	<i>n</i>	<i>Percentage</i>
Couples group	11	91.7	1	8.3
Gender-specific group	13	92.9	1	7.1

NOTE: $\chi^2 = 0.0129(1)$; $p = .91$; Cramér's $V = .0223$.

subject who had been assigned to the gender-specific group who had returned to court for adjudication for further episodes of violence.

Safety

Six instances of ongoing physical or emotional abuse were reported to the facilitators or recorded on the TPWSs. Of these, 2 involved couples assigned to the couples group intervention and 4 involved couples assigned to the gender-specific intervention. Consequently, in this study there was no evidence to support the belief that women who received treatment conjointly were placed in any further jeopardy than those who attended classes separate from their spouses.

DISCUSSION AND APPLICATIONS TO SOCIAL WORK PRACTICE

The foregoing results indicate that for the largest number of subjects (those with no history of alcohol abuse) neither approach to intervention appears to be more effective than the other. However, for subjects with a history of alcohol abuse the couples group intervention appeared to be more effective than the gender-specific intervention in reducing the level of violence within the marital relationship. All of the subjects known to have a history of alcohol abuse were involved in a court-monitored Antabuse program throughout this treatment program. Thus it can be assumed they were alcohol-free. To what, then, can the differences be attributed? One of the major advantages to using the group format in treating substance abusers is that the effect of confrontation is maximized. When one individual is confronted by the facilitator, others in the group who struggle with the same issue are indirectly confronted. In addition, the ability to minimize or project responsibility is limited due to the

fact that the other members are quick to point out these projections. The issue of group confrontation is critical to the findings of the current study. Furthermore, it is possible that the couples treatment had the added benefit of addressing relationship problems caused by alcohol abuse. It is important to note that further studies are critical to evaluating the importance of this finding.

Flanzer (1993) suggested that "alcohol may serve as a rationalization for violence, allowing the perpetrator to avoid taking responsibility for his or her actions" (p. 178). This abrogation of responsibility is based on the belief that the abuser cannot remember actions that took place while he or she was under the influence of alcohol and thus was not responsible for what might have occurred. For those abusers attending classes along with their partner, the opportunity to deny, minimize, or project responsibility onto the victims may be reduced because they are likely to be confronted in the presence of other couples by either their spouse, the facilitator, or other group members.

Extreme caution should be exercised in generalizing the results of this study. The sample was limited to intact couples who had not yet experienced relatively severe forms of abuse and who had indicated a desire to remain in their current relationship. It was comprised of predominantly lower- to lower-middle-class individuals who have come into contact with the legal system. As an incentive to participate, subjects were provided services at a substantially reduced fee rate, whereas their partners were provided services pro bono. Prior studies (Gelles & Cornell, 1990; Labell, 1979) have documented the difficulties of making broad generalizations to the population as a whole based on studies of this type. Because all of the subjects in this study were court-mandated into treatment, the results achieved might actually be due to the surveillance resulting from probation. Future investigators might want to include both court-referred and voluntary clientele to expand the generalizability of the findings.

This study is further limited by the fact that, as in the case for the vast majority of studies involving violence, there was no nontreatment control group. The results achieved could actually be the product of maturation, the natural remission of abuse, or the decision to stay together, and so on. In addition, the relatively small sample size raises the possibility of failing to find a significant difference when one is actually present. Future efforts at replication should ensure that sufficient subjects are involved to enhance the ability to generalize the findings.

Lipchik, Sirles, and Kubicki (in press) argued that treatment options present a dilemma between the sociopolitical philosophy that considers domestic violence as a male tactic to control women and the reality of

multifaceted acts of violence. Social workers need to be flexible in their treatment approaches and establish the best treatment plan based on the individual needs of the client rather than being forced to adhere to a sociopolitical philosophy. In some instances, the safety of the victim is so imperiled that her safety and the dissolution of the relationship are the most appropriate treatment objectives. Other couples, however, might benefit more from a couples group approach that does not attempt to dissolve the relationship.

Although clinicians and researchers are seeking new interventions to eliminate the incidence of spouse abuse, some advocates lobby legislators to restrict the use of state funds for domestic violence treatment programs to those programs that treat only male batterers and that specifically exclude couples and family counseling options. Research and development of safe and effective alternative interventions, such as couples and family treatment, are not being invited. In fact, some are discouraged categorically (Brandl, 1990).

Currently, there exists a strong trend toward individual states funding only those programs that adhere to the philosophy of treating only the batterer. Colorado began this trend in 1988 by establishing standards that prescribed group programs for batterers as the primary treatment method and forbade the use of couples groups. This emphasis on batterer-only groups has spread to Connecticut, New York, Texas, and Minnesota. Pennsylvania is currently working on standards to control which domestic violence programs receive court-ordered clients (Lipchik et al., in press).

Because 50% to 75% of women who have been involved in relationships that are violent decide to remain in them despite the best efforts of police, prosecutors, shelters, and advocates (Feazell, Mayers, & Deschner, 1984; Purdy & Nickle, 1981), it would seem reasonable for legislators not to prohibit funding of programs that will provide services to this population. Indeed, based on the findings reported in the current study, the couples group format was at least as effective as the batterer-only groups for the majority of subjects, and superior for those with an alcohol-related problem. In addition, decision makers need to be made aware that there remains a dearth of research supporting the notion that women attending couples groups are in any more jeopardy than those who receive counseling apart from their spouses or, for that matter, who receive no treatment at all.

Future research is needed to replicate the findings of this study. Investigators might consider using a no-treatment control group in addition to the couples and gender-specific groups. With a no-treatment control group, we can ascertain whether pretest to posttest gains for both groups, as in the results of the current study, indicate that both intervention approaches are effective.

REFERENCES

- American College of Obstetricians and Gynecologists. (1989). *The battered woman* (Technical Bulletin No. 124). Washington, DC: Author.
- Brandl, B. (1990). *Programs for batterers: A discussion paper*. Madison, WI: Department of Health and Social Services.
- Caesar, P. L., & Hamberger, L. K. (Eds.). (1989). *Treating men who batter: Theory, practice, and programs*. New York: Springer.
- DAP Women's Therapy Team. (1991). *Women's treatment handbook*. Minneapolis, MN: Domestic Abuse Project.
- Deschner, J. P. (1984). *The hitting habit: Anger control of battering couples*. New York: Free Press.
- Edleson, J. L., & Brygger, M. P. (1986). Gender differences in reporting of battering incidences. *Family Relations*, 35, 377-382.
- Edleson, J. L., & Syers, M. (1990). Relative effectiveness of group treatments for men who batter. *Social Work Research and Abstracts*, 26(2), 10-17.
- Edleson, J. L., & Tolman, R. M. (1992). *Intervention for men who batter*. Newbury Park, CA: Sage.
- Epstein, N. B., Baldwin, L. M., & Bishop, D. S. (1983). The McMaster family assessment device. *Journal of Marital and Family Therapy*, 9, 171-180.
- Feazell, C. S., Mayers, R. S., & Deschner, J. (1984). Services for men who batter: Implications for programs and policies. *Family Relations*, 33, 217-223.
- Flanzer, J. (1993). Alcohol and other drugs are key causal agents of violence. In R. Gelles & D. Loseke (Eds.), *Current controversies on family violence* (pp. 171-181). Newbury Park, CA: Sage.
- Ganley, A. L. (1989). Integrating feminist and social learning analyses of aggression: Creating multiple models for intervention with men who batter. In P. L. Caesar & L. K. Hamberger (Eds.), *Treating men who batter: Theory, practice, and programs* (pp. 196-235). New York: Springer.
- Geller, A. J. (1982). Conjoint therapy: Staff training and treatment of the abuser and the abused. In M. Roy (Ed.), *The abusive partner* (pp. 198-215). New York: Van Nostrand Reinhold.
- Gelles, R. J., & Cornell, C. P. (1990). *Intimate violence in families* (2nd ed.). Newbury Park, CA: Sage.
- Gurman, A. S., & Kniskern, D. P. (1978). Research on marital and family therapy: Process, perspective and prospect. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (2nd ed., pp. 817-901). New York: Wiley.
- Gurman, A. S., Kniskern, D. P., & Pinsof, W. M. (1986). Research on the process and outcome of marital and family therapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed., pp. 565-624). New York: Wiley.
- Harris, J. (1986). Counseling violent couples using Walker's model. *Psychotherapy*, 23, 613-621.
- Jouriles, E. N., & O'Leary, D. K. (1985). Interspousal reliability of reports of marital violence. *Journal of Consulting and Clinical Psychology*, 53, 419-421.
- Kilpatrick, D. (1992). *Rape in America: A report to the nation*. Washington, DC: The National Victim Center and the Crime Victims Research and Treatment Center at the Medical University of South Carolina.
- Labell, L. S. (1979). A sociological study of battered women and their mates. *Victimology: An International Journal*, 4, 257-267.

- Lipchik, E., Sirles, E. A., & Kubicki, A. D. (1992). *Multifaceted approaches to spouse abuse treatment*. In R. Geffner, S. B. Sorenson, & P. K. Lundberg (Eds.), *Current research and interventions in domestic violence*. Binghamton, NY: Hayworth.
- Margolin, G. (1979). Conjoint marital therapy to enhance anger management and reduce spouse abuse. *American Journal of Family Therapy*, 7, 13-24.
- Miller, I. W., Epstein, N. B., Bishop, D. S., & Keitner, G. I. (1985). The McMaster family assessment device: Reliability and validity. *Journal of Marital and Family Therapy*, 11, 345-356.
- Neidig, P. H. (1984). Women's shelters, men's collectives and other issues in the field of spouse abuse. *Victimology*, 9, 464-476.
- Neidig, P. H. (1986). *A modified conflict tactics scale* (unpublished scale). Stony Brook, NY: State University of New York at Stony Brook.
- Neidig, P. H. (1992). *Treatment project weekly summary* (unpublished scale). Stony Brook, NY: State University of New York at Stony Brook.
- Neidig, P. H., & Friedman, D. H. (1984). *Spouse abuse: A treatment program for couples*. Champaign, IL: Research Press.
- Ohlsen, M. M. (1979). *Marriage counseling in groups*. Champaign, IL: Research Press.
- Pan, H. S., Neidig, P. H., & O'Leary, K. D. (1994). Male-female and aggressor-victim differences in the factor structure of the modified conflict tactics scale. *Journal of Interpersonal Violence*, 9, 366-382.
- Ptacek, J. (1988). Why do men batter their wives? In K. Yllo & M. Bograd (Eds.), *Feminist perspectives on wife abuse* (pp. 133-157). Newbury Park, CA: Sage.
- Purdy, F., & Nickle, N. (1981). Practice principles for working with groups of men who batter. *Social Work with Groups*, 4, 111-122.
- Rosenbaum, A., & O'Leary, K. D. (1986). The treatment of marital violence. In N. S. Jacobsen & S. Gurman (Eds.), *Clinical handbook of marital therapy* (pp. 385-405). New York: Guilford.
- Rusinoff, J. S. (1990). *Men's treatment handbook* (2nd ed.). Minneapolis, MN: Domestic Abuse Project.
- Sherman, L., & Berk, R. (1984). The specific deterrent effects of arrest from domestic violence. *American Sociological Review*, 49, 261-272.
- Snyder, D. K. (1981). *Marital satisfaction inventory (MSI) manual*. Los Angeles: Western Psychological Services.
- Sonkin, D. J., Martin, D., & Walker, L. (1985). *The male batterer: A treatment approach*. New York: Springer.
- Stark, E., Flitcraft, S. E., & Frazier, W. (1979). Medicine and patriarchal violence: The social construct of a private event. *International Journal of Health Services*, 9, 461-498.
- Stevens, J. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Straus, M. A., & Gelles, R. J. (1986). Societal change and change in family violence from 1975 to 1985 as revealed by two national surveys. *Journal of Marriage and the Family*, 48, 465-479.
- Straus, M. A., & Gelles, R. J. (1988). How violent are American families? Estimates from the national family violence resurvey and other studies. In G. T. Hotaling, D. Finkelhor, J. T. Kirkpatrick, & M. A. Straus (Eds.), *Family abuse and its consequences* (pp. 14-36). Newbury Park, CA: Sage.
- Straus, M. A., & Gelles, R. J. (Eds.). (1990). *Physical violence in American families*. New Brunswick, NJ: Transaction.

- Szinovacz, M. E. (1983). Using couple data as a methodological tool: The case of marital violence. *Journal of Marriage and the Family*, 45, 633-644.
- Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics* (2nd ed.). New York: Harper Collins.

The San Diego Navy Experiment: An Assessment of Interventions for Men Who Assault Their Wives

Franklyn W. Dunford
University of Colorado at Boulder

Three different 12-month interventions for servicemen who had been substantiated as having physically assaulted their wives were used and the outcomes examined. The 861 couples of the study were randomly assigned to 4 groups: a men's group, a conjoint group, a rigorously monitored group, and a control group. Cognitive-behavioral interventions were implemented for the men's and conjoint groups, and outcome data were gathered from male perpetrators and female victims at roughly 6-month intervals over the approximately 18-month experimental period. Data analyses revealed nonsignificant differences between the experimental groups over a variety of outcome measures.

Reviews of treatment services for men who abuse their wives or cohabitant partners reveal that the majority of group interventions for men who batter have been based on a cognitive-behavioral model (Eisikovits & Edleson, 1989; Hamberger & Hastings, 1993; Rosenfeld, 1992; Tolman & Bennett, 1990), that evaluations have routinely failed to use rigorous experimental designs (Chalk & King, 1998; Fagan, 1996), and that little evidence exists that the prevailing interventions for men who batter are efficacious (Crowell & Burgess, 1996; Healey, Smith, & O'Sullivan, 1998). The ability to rule out alternative explanations for what appear to be positive findings regarding the cessation of spouse or partner abuse has been notably absent in almost all evaluations of domestic violence programs. The net effect is a lack of reliable information about how to best treat men who abuse their wives or cohabitant partners (Boruch, 1994; Fagan, 1996).

A literature search conducted before the San Diego Navy Experiment was initiated in 1991 found only one evaluation of an intervention conducted in a military setting for wife abuse. In that evaluation of the Domestic Conflict Containment Program (Neidig, 1985) implemented with Marine Corps men, Neidig (1986) indicated that postintervention measures showed significant and positive changes on the Dyadic Adjustment Scale subscales, con-

sensus and cohesion measures, and the Norwicki Strickland Locus of Control Scale. In 1992, Mollerstrom, Patchner, and Milner referred to an evaluation of programs for batterers in the Air Force, but no published results have been found in the literature for that or for any other evaluation of military batterer treatment programs. Neither of these evaluations used an experimental research design.

A much larger number of evaluations have been conducted on programs for male batterers in civilian settings, but with few exceptions (Davis & Taylor, 1997; Feder, 1998; Palmer, Brown, & Barrera, 1992) the evaluations were no more rigorous than those found for the military. According to Rosenfeld (1992), as of 1992 only three research projects for spouse or partner abuse had randomly assigned participants to different treatment conditions, and only one of those used what could be conceptualized as a control group.

The present study, the San Diego Navy Experiment, was designed to experimentally evaluate the effectiveness of cognitive-behavioral interventions implemented in different treatment settings for men who batter. This intervention model was selected for evaluation for several reasons. As noted earlier, it was the most prevalent group intervention for men who batter their cohabitant partners used at the time the San Diego project was initiated (Tolman & Bennett, 1990). In addition, a relatively large number of uncontrolled evaluations suggested, at that time, that interventions based on the cognitive-behavioral model effectively addressed the problem of the continued abuse of cohabitant partners. Finally, it was the treatment model most frequently used by the Navy. (For a brief description of the cognitive-behavioral model, see Saunders, 1999.)

The decision to evaluate the effectiveness of a cognitive-behavioral intervention in both men's and conjoint groups was based primarily on existing Navy practice. In addition, at the Navy's request, the study evaluated a third intervention that used a "stake in conformity" strategy (Toby, 1957). The objective was to determine if male perpetrators held accountable for their abusiveness toward their wives, using systematized and official monitoring procedures, would stop the continued abuse. This intervention was called *rigorous monitoring*.

The grant (2 RO1MH4508) supporting this research was awarded and administered by the National Institute of Mental Health (NIMH). All funds were provided by the U.S. Navy through an interagency agreement with NIMH. The contents of this report reflect the views of the author and not necessarily the position of either the U.S. Navy or NIMH.

I am indebted to the officials of the U.S. Navy, who, contrary to common research practice in this area, supported a strong experimental research design to determine if the interventions provided to Navy men who abuse their wives were effective. I am indebted, as well, to Sandra Rosswork, U.S. Navy, and James Breiling, NIMH, for the major contributions they made to this experiment. I thank the staff of the Family Advocacy Center, San Diego Naval Base, and the Colorado University research team for their loyal support, as well as the scholars who so thoroughly and thoughtfully reviewed an earlier version of this article.

Correspondence concerning this article should be addressed to Franklyn W. Dunford, Institute of Behavioral Science, University of Colorado, Boulder, Colorado 80309.

Method

Participants

The experimental sample consisted of 861 married U.S. Navy couples in which active-duty husbands were substantiated as having physically assaulted their wives. The sample was young (mean age = 27 years, mode = 24 years), newly married (mean length of marriage = 47 months, mode = 24 months), with children (83%, mean number of children = 1.7), moderately educated (mean number of school years = 12.6), mostly White (men = 48%, women = 40%) or Black (men = 35%, women = 28%), and low in rank (mean = Petty Officer Third and Fourth Class, the equivalent of Private First Class and Corporal, respectively, in the Army), with low to moderate incomes ($M = \$2,594$ total family income per month before taxes).

Interventions

The study participants were randomly assigned to a men's group, conjoint group, rigorous monitoring group, or control group.

Men's group. The men's group, which used a cognitive-behavioral model of change, met weekly for 6 months and then monthly for another 6 months, for a total 1-year treatment period. The curriculum for this intervention was developed by Daniel Saunders and David Wexler based on their work in domestic violence (Saunders, 1996; Wexler, 1999). The weekly meetings included both didactic and process activities.

In the didactic portion of the weekly sessions, group leaders addressed perpetrator attitudes and values regarding women and violence toward women and taught the men a wide variety of skills thought to be important to the successful elimination of the continued abuse of women (e.g., cognitive restructuring, empathy enhancement, communication skills, anger modification, and jealousy). All of the skills taught were derived from the cognitive-behavioral model. Each session involved a set of tasks that group leaders were obliged to complete. Session 2, for example, specified the following eight tasks: review weekly check-in form, review the nine commandments (statements ranging from "We are all 100% responsible for our behavior" to "Counselors and case managers cannot make people change—they can only set the stage for change to occur"), explain the need for a responsibility plan, identify warning signals, explain "time-out," review each step of a responsibility plan, guide all members in developing individual responsibility plans, and assign time-out homework. The process portion of the weekly sessions involved dealing with issues raised in the didactic segments of the sessions as well as with other issues that emerged. The six monthly sessions that followed consisted of review and process activities.

Conjoint group. Despite the controversy associated with treating victims and perpetrators of spouse abuse together (Edleson & Tolman, 1992), the Navy was interested in assessing a conjoint group treatment approach because of the long-standing use of couples-group approaches. The conjoint group was organized in much the same way as the men's group, with 26 weekly sessions that included both didactic and process activities followed by 6 monthly sessions. The curriculum for this intervention was also based on the cognitive-behavioral model and was developed by Geffner (2000). The interventions were similar to those used in the men's group, with the major difference being that the presence of wives was expected to alter the dynamics of the conjoint group interventions. The presence of wives in the conjoint group was expected, for example, to promote realistic and personalized role-playing, reduce "women bashing," personalize violence (including intimidation and emotional abuse), and enhance empathy in ways more poignant than were evident in the men's group. In addition, the ability of wives to witness authority figures confronting the offensive and oppressive nature of spouse abuse, as well as address constructive ways to deal with conflict, were proposed as sources of empowerment and confidence not available to women whose husbands were assigned to the other interventions. These kinds of contextual differ-

ences were expected to increase the effectiveness of the cognitive-behavioral treatment model. As in the men's group, the 6 monthly sessions consisted of review and process activities.¹

Rigorous monitoring group. The rigorous monitoring intervention represents the formalization and systematization of the Navy's attempts to hold perpetrators accountable for their abusiveness toward their wives. The goal of the interventions was to inhibit continued abuse by making new instances of abuse more visible to those having control over the lives of service members (i.e., their commanding officers). Perpetrators were seen monthly for individual counseling for 12 months by a case manager at the Family Advocacy Center (FAC), the Navy agency responsible for the treatment of men who abuse their wives. Every 6 weeks a record search was completed to determine if perpetrators had been arrested or referred to court anywhere in San Diego County. Wives were called monthly and asked about new instances of abuse. They were advised during the calls that they were not obliged to reveal anything to the FAC representative about their husbands' behavior if doing so would place them in jeopardy (which may have reduced the effectiveness of this intervention). Case managers delivered individual counseling at each session as they thought appropriate. Case managers were required, as a part of the rigorous monitoring treatment intervention, to inform clients that their behavior was being monitored and that their commanding officers would be advised monthly of all new instances of abuse. At the end of each treatment session, case managers sent progress reports to perpetrators and their commanding officers specifying the presence or absence of new instances of abuse. In this manner, an attempt was made to create a "fishbowl" effect for the male participants of the rigorous monitoring group.

Control group. Men assigned to the control group were to receive no FAC treatment. However, their wives, like the wives of the men assigned to the other treatment groups, did receive preliminary assistance called *stabilization and safety planning*. That is, the FAC contacted victimized wives as soon as possible after the presenting incident to ensure that they were not in immediate danger of continued abuse. Once their safety was assured, the FAC provided them with safety planning information.

Random Assignment

Sample selection took 46 months. The majority of the sample came from referrals from the ships or stations to which servicemen were assigned, the Family Service Centers (Navy agencies tasked to provide support and educational services to Navy families), and Navy medical facilities. Once a serviceman was determined to be eligible for the research, the case was given to the research staff and subsequently entered into a computer programmed to randomize cases to one of the four experimental treatments. For a serviceman to be deemed eligible for the research, the Navy Case Review Committee had to substantiate that he had physically assaulted his wife, that divorce procedures were not officially in process, that he had more than 6 months left to serve in the area, that he was not alcoholically impaired, and that he was devoid of significant pathology (i.e., he did not present with active psychosis, antisocial personality disorder, pathological jealousy, or suicidal ideation). Servicemen suspected of alcoholic impairment were required to go through an official assessment for alcoholism called a *Counseling and Assistance Center (CAAC) screen*. These screens resulted in three outcomes: Level I (no ongoing treatment required), Level II (4 weeks of outpatient counseling), and Level III (4 to 6 weeks of in-patient treatment conducted in a Navy hospital). Twenty percent of the eligible cases were referred for CAAC screens. Fifty-three percent of those cases were subsequently returned to the FAC for treatment when identified as nonalcoholic or when they had completed treatment for their alcoholism.

¹ The details of the curricula for the men's and conjoint groups are too lengthy to describe in this report. Manuals that specify the content and stated objectives for all of the group sessions for each of these approaches (U.S. Navy, 1993a, 1993b) are available from Franklyn W. Dunford.

Measures

Two major types of measures were used in the analyses for this report: demographic and outcome assessments. The demographic measures assessed such standard variables as ethnicity, income, education, age, rank, and family size. Four types of outcome measures were used. A self-reported episodic measure assessed the number of incidents or episodes in which a victim or perpetrator reported being abused across three different levels of abuse. Female victims were asked, for example, to indicate how many incidents had occurred in which they (a) felt like they were in danger of being hurt by their husbands; (b) they were pushed, hit, or had hands laid on them, or were beaten up by their husbands; and (c) were physically injured by their husbands (e.g., knocked down, bruised, scratched, cut, choked, had bones broken, had eyes or teeth injured, or were still hurting the next day).

The second outcome measure, the Modified Conflict Tactics Scale (MCTS; Straus, 1979), focused on types of abusive behaviors as reported by respondents. The items of the MCTS differed from the original CTS in two ways. The "double-barreled" items of the original scale were divided into separate questions, and a few items were added to gather additional information about what can best be described as psychological abuse. For each of the 42 items of the MCTS, respondents were asked to indicate the frequency with which they engaged in or were subject to a specific behavior. For scoring, the items were combined into eight subscales, two of which (All Violence and Severe Violence) were patterned after traditional CTS measures (Straus, 1979). The following is a brief description of the subscales, with their attendant reliability (standardized item alpha) levels for both men and women reports: Passive Abuse consisted of such items as "Treated her as if she was inferior" and "Ignored her requests or feelings" (for women, $\alpha = .89$; for men, $\alpha = .88$); Control Abuse consisted of items like "Controlled money against her will" and "Refused to let her get or keep a job" (for women, $\alpha = .77$; for men, $\alpha = .38$); Menacing Abuse consisted of such items as "Looked at her in ways that she knew he meant to hurt her" and "Destroyed furniture, walls, objects and things because he was mad at her" (for women, $\alpha = .70$; for men, $\alpha = .61$); Misdemeanor Abuse consisted of items like "Grabbed, restrained, held on to her against her will" and "Pushed shoved or slapped her" (for women, $\alpha = .78$; for men, $\alpha = .75$); Felony Abuse consisted of such items as "Hit her with a fist" and "Threatened her with a knife or gun" (for women, $\alpha = .81$; for men, $\alpha = .73$); Sexual Abuse consisted of the items "Verbally pressured her to have sex" and "Forced her to have sex against her will" (for women, $\alpha = .50$, for men, $\alpha = .87$). The All Violence (for women, $\alpha = .86$; for men, $\alpha = .80$) and Severe Violence (for women, $\alpha = .84$; for men, $\alpha = .76$) subscales were aggregates of items from the Misdemeanor Abuse and Felony Abuse subscales. In all of these measures, high scores indicated high levels of abuse. The measures were developed as theoretical, rather than empirically derived, constructs.

The third outcome measure consisted of official police and court records for all respondents (both victims and perpetrators) living within the boundaries of San Diego County. The fourth outcome measure focused on the date of the first instance in which a repeat case of spouse assault occurred as indicated by both official arrest records and victim reports of new physical injuries.

Interviews With Victims and Perpetrators

The research design called for the University of Colorado research team (not FAC/Navy staff) to interview victims and perpetrators separately, typically in their homes, four times over the course of the experiment at approximately 6-month intervals. A baseline measure was taken before treatment began, a second interview was conducted at the conclusion of the first 6 months of treatment, and two more interviews were conducted at subsequent 6-month intervals. The interviews were designed to measure the outcome effects of the interventions. Extensive provisions were made to ensure that respondents knew that the interviewers were employees of

the University of Colorado, that the information they shared with the interviewers would be sent directly to the University of Colorado, and that neither the Navy nor spouses would ever have access to individual responses to the interviews. For example, participants were given a letter from the admiral of the San Diego Naval Base assuring them that the information given to university interviewers would be sent directly to the University of Colorado, that the Navy would never have access to it, and that their answers could in no way adversely affect them. Further, respondents were given a certificate of confidentiality issued by the Department of Health and Human Services that protects survey data from subpoena. Each participant also received an incentive payment of \$50 for a completed interview.

Not all the participants agreed to be interviewed by the research staff. The cumulative completion rate was as follows: first interview, 86%; second interview, 82%; third interview, 78%; and fourth interview, 75%. Refusals accounted for the majority of cases that did not complete the first interview, and the inability to find or contact people accounted for the majority of later losses (individuals and couples moved away from the area; were no longer in the Navy; were no longer living together [separated or divorced]; or, as was often the case, took steps to avoid being found). Comparing cases lost and not lost by the fourth interview, using the baseline measures of episodic and MCTS abuse, a tendency was found for those who failed to be interviewed to be slightly more involved in prior abusiveness as reported by wives than those who agreed to be interviewed, although only 2 of the 11 comparisons between the groups (Misdemeanor Abuse and Sexual Abuse) yielded results that were statistically significant. No statistically significant differences or trends were found for the same comparisons for men.

Results

Equivalency

Comparisons of the experimental groups on the prevalence and frequency of episodic abuse during the 6-month period prior to the presenting incident revealed no statistically significant differences. Similarly, comparisons using the MCTS subscales for abuses occurring during the 6-month period prior to the presenting incident and for abuses occurring during the presenting incident showed fewer differences than would be expected to occur by chance. Furthermore, nonsignificant differences were found between the groups for a set of mediator/predictor variables used in the research; for the prevalence or frequency of prior arrests for spouse abuse; and for the demographic variables of ethnicity, age, education, income, rank, injury in prior relationships, witnessing of parental violence, number of children in the family, victim employment status, alcohol use, and drunkenness. The results suggest that random assignment was effective.

Treatment Integrity

The curricula for each of the 26 sessions for the men's groups, conjoint groups, and the rigorous monitoring sessions specified the tasks that were to be completed for each session. Each group session lasted 1.5 hr and was conducted by a male cotherapist and a female cotherapist. To ensure uniformity of services over time and across therapists, all of the group leaders conducted both men's and conjoint groups, thus controlling, in part, the effects of therapist characteristics. The rigorous monitoring sessions were typically 1-hr sessions conducted by individual case managers. All of the sessions (men's, conjoint, and rigorous monitoring) were audiotaped.

Sample audiotapes for each therapist and case manager were evaluated monthly over the course of the experiment. Research staff listened to each of the tapes in its entirety to determine the extent to which appointed tasks were completed (adherence) and to evaluate the performance skills of the group leaders. Adherence to treatment protocols was relatively high for all three treatments (for the men's group, 78% thoroughly addressed, 7% moderately addressed, 5% introduced, and 3% mentioned; for the conjoint group, 72% thoroughly addressed, 9% moderately addressed, 5% introduced, and 3% mentioned; for the rigorous monitoring group, 82% thoroughly addressed and 7% introduced). Findings from the reviews of the audiotapes were reported to the group leaders and case managers as the data were collected. The between-coders correlation (Pearson's r) was .93.

With regard to performance skills, the audiotapes were evaluated to determine, for example, how well the group leaders integrated process and didactic material, used appropriate confrontation skills with clients, and provided suitable feedback. Less than 5% of the group therapists were rated below an "at-standard" level of performance for any of the performance dimensions assessed over the course of the study.

Dropouts

At the conclusion of prescribed treatment periods, the Navy's FAC formally closes all cases for which treatment has been assigned. FAC records indicated that 71% of the cases involved in the analyses for the 1-year follow-up period were closed as successfully completing treatment. Another 15% of the men were discharged from the Navy during treatment; and although not dropouts in the traditional sense, they did not complete treatment. The remainder of the cases (14%) were not closed as completing treatment because the men were transferred out of the area, recidivated and received another treatment, or failed to attend the required treatment sessions.

The proportion of men who failed to complete treatment was larger than expected. Discharges from the Navy, transfers, deployments, and a lack of support from commanding officers had negative effects on completion rates. Although space limitation prevents the presentation of detailed findings in this report, one finding is especially noteworthy. Analyses of attendance data indicated that the continued abuse of wives was inconsistently and only limitedly affected by the amount of treatment received, a finding that is consistent with data reported elsewhere (Palmer et al., 1992).

Crossovers

A crossover occurred when a case was randomly assigned to one group but received an alternative treatment. Twenty-three cases (3% of the sample) were crossovers. Crossovers occurred for a variety of reasons, the most common being deployment schedules, which required men to be in and out of port on a regular basis over relatively long periods of time. Such cases were often moved from a group treatment approach, which required meeting at preset and regularly scheduled times, to individualized treatment, where services could be delivered on an ad hoc basis. Also, the local courts would occasionally require that a case referred to the FAC and assigned to the experiment be reassigned to a specific type of

treatment. Unique and relatively uncommon circumstances were responsible for other types of crossovers. Among those circumstances were mistakes in the assignment process, serious sick or terminally ill clients, and reassignment to groups outside of the experiment following the emergence of new information. When analyses were conducted on the basis of treatments as delivered or when the 23 crossover cases were eliminated from the analyses, outcome findings did not differ from those obtained from analyses based on treatment as assigned. All of the analyses presented in this article are based on treatment as assigned, which is the protocol suggested by Weinstein and Levin (1989) and used by others (Davis & Taylor, 1997).

Conjoint Group Attendance

Notwithstanding efforts to increase the attendance of wives in the conjoint group, including the provision of funds for child care and the use of victim advocates, the average number of wives attending the conjoint group sessions was relatively low. The average ratio of attendance of women to men was 2:5. Given the role that women were hypothesized as playing in the conjoint group and the variability in the attendance ratios across time (from a low of 1:10 to a high of 7:10), analyses were conducted to assess the effects that the attendance of varying numbers of women may have had on outcome.

Three different measures of female victim attendance at conjoint group sessions were assessed: (a) the average number of women attending the sessions; (b) the number of sessions attended by men when there were 1, 2, 3, or more women in attendance; and (c) the average number of sessions in which a woman attended with her husband. Analyses were completed using the total sample, restricting the sample to men who attended 20 or more sessions, and restricting the sample to men attending 10 or more sessions. When simple correlations were calculated between the various women's attendance variables and the frequency of episodic abuse reported at the fourth interview for all versions of the samples, no statistically significant correlations were found. Similarly, when multivariate regression models were used that added prior abuse, living together, and men's attendance records as independent variables, no statistically significant relationships between victim attendance and continued abuse were found. These data suggest that victim attendance, as operationalized and measured here, did not have the impact envisioned for the conjoint group intervention, although the findings might have been different had the attendance ratio of women to men been higher.

Victim- and Perpetrator-Reported Outcome

Results of comparisons of the four experimental groups on all of the outcome variables assessed at the third and fourth interviews (combined) using a variety of prevalence and frequency measures are shown in Tables 1–4. Third- and fourth-interview outcome data were combined for several reasons. First, second-interview data (collected at the end of the first 6 months of treatment) could not be used as a primary measure for treatment effects because of the time-order confound between the number of sessions attended and participant-reported recidivism during the second interview period. Since the dates of all new abuses were not recorded, new instances of abuse could have occurred early in the second mea-

Table 3
The Frequency of Continued Abuse by Treatment Group at 1-Year Follow-Up Using the Modified Conflict Tactics Scale (MCTS)

MCTS subscale	Men's group ^a	Conjoint group ^b	RM group ^c	Control group ^d	F ^e	p
Wives' reports						
Passive Abuse	144.97	197.85	141.21	159.67	0.818	.48
Control Abuse	54.68	48.71	30.64	33.42	1.184	.31
Menacing Abuse	13.44	16.68	29.33	13.69	0.973	.40
Misdemeanor Abuse	8.88	6.45	20.37	6.34	0.904	.44
Felony Abuse	6.49	1.60	4.11	.64	0.893	.44
Sexual Abuse	3.49	4.15	3.87	1.77	0.333	.80
All Violence	9.48	4.79	13.29	3.51	0.746	.52
Severe Violence	6.20	1.82	3.74	1.10	0.657	.58
Husbands' reports						
Passive Abuse	15.87	27.04	52.90	33.07	1.997	.11
Control Abuse	2.91	9.42	6.81	9.45	1.300	.27
Menacing Abuse	1.08	2.16	2.83	1.34	0.770	.51
Misdemeanor Abuse	0.60	0.86	2.96	1.01	1.030	.38
Felony Abuse	0.04	0.08	0.11	0.01	0.788	.50
Sexual Abuse	0.00	0.04	0.92	0.04	1.223	.30
All Violence	0.35	0.62	2.60	0.40	1.050	.37
Severe Violence	0.08	0.21	0.19	0.06	0.880	.45

Note. RM = rigorous monitoring.

^a *n* = 162 for wives, 160 for husbands. ^b *n* = 158 for wives, 146 for husbands. ^c *n* = 155 for wives, 169 for husbands. ^d *n* = 145 for wives, 144 for husbands. ^e For wives' reports, *df* = 3, 616; for husbands' reports, *df* = 3, 615.

surement period before many of the weekly treatments had been delivered, midway through the measurement period, or late in the measurement period after most of the weekly treatment had been delivered, obscuring the relationship between the amount of treatment received and new instances of abuse. To avoid the confound, the decision was made to measure outcome after the weekly sessions had ended. Second, because the men's and conjoint groups were the two major interventions of the experiment, and because the effects of treatment for these two groups were expected to be greatest at the end of the first 6 months of weekly sessions, third- and fourth-interview data were expected to most accurately capture the outcome data for these two groups. Third, when experimental analyses were conducted separately for data obtained from second, third, and fourth interviews, the results produced the same findings as those found when third- and fourth-interview data were combined, thus supporting the use of the combined data and facilitating a more parsimonious presentation

of the findings. Thus, all of the outcome measures used in this report involve the 1-year follow-up period following the first 6 months of treatment.

Table 1 shows victim reports of episodic spouse abuse, using the combined third- and fourth-interview outcome data. No statistically significant differences ($p \leq .05$) were found between the four experimental groups for the prevalence of continued abuse using these measures. The frequency with which women were victimized is also reported in Table 1. Although the results were not statistically significant, women from the rigorous monitoring group reported that their husbands recidivated at much higher frequencies on all three of the episodic measures than did the husbands from the other experimental groups. These high numbers are primarily the result of extreme scores reported by a few women in the rigorous monitoring group (1 of the women, for example, reported over 1,000 incidents of abuse). When a log transformation of the frequency of reoffending was used to reduce the influence

Table 4
Arrest Recidivism by Treatment Group at 1-Year Follow-Up

No. of arrests	Men's group (<i>n</i> = 168)		Conjoint group (<i>n</i> = 153)		RM group (<i>n</i> = 173)		Control group (<i>n</i> = 150)		Total (<i>N</i> = 644)	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
0	162	96	149	97	163	94	144	96	618	96
1	6	4	4	3	10	6	6	4	26	4
Frequency	0.036		0.026		0.058		0.040		0.040	

Note. $\chi^2(3, N = 644) = 2.251, p = .52$. RM = rigorous monitoring.

of the extreme scores, the differences between the groups disappeared.

Table 2 summarizes male perpetrator reports of the abuse of their wives. No statistically significant differences between the groups were found for any of the comparisons for the prevalence or frequency measures. Differences were less when log transformations were used to eliminate the influence of extreme scores.

Comparisons of data for the episodic measures, for both victim and perpetrator reports of new abuse (see Tables 1 and 2), produced no evidence that membership in any of the three experimental treatment groups was any more effective in reducing continued abuse than was membership in the control group.

Comparisons of the frequency of continued abusive behaviors for the different groups using the MCTS are shown in Table 3. None of the differences between groups for either the women's or the men's reports of the continued abuse of wives were found to be statistically significant. The apparent substantive differences between the means of some of the subscales of Table 3 were due, once again, to the extreme scores reported by a few respondents.

To further assess the effects of treatment, the men's and conjoint groups were collapsed to form one treatment group to be compared with the control group using the same comparative techniques and outcome variables represented in Tables 1–3. Similar analyses were completed by collapsing the men's, conjoint, and rigorous monitoring groups into one group and comparing it with the control group. None of these assessments altered any of the findings presented earlier. Further, pairwise *t* tests conducted on all possible combinations of treatment groups showed no statistically significant differences for any of the comparisons.

Abuse reported by victimized wives/partners is almost always found in the literature to be greater than that reported by perpetrating husbands/partners (Szinovacz, 1983), a finding clearly reflected in Tables 1–3. To address this disparity, victim and perpetrator reports of abuse were combined and assessed. The results did not alter any of the findings.

Official Outcome Data

The data in Table 4 represent all arrests in which the same perpetrating husband revictimized the same victimized wife during the 12-month period following the first 6 months of treatment. No statistically significant difference was found between the groups for the prevalence of new arrests.

Survival Analyses

Survival analyses were applied to data from two sources to determine if time to recidivism varied for the different experimental groups. First, wives were asked to indicate the date of the first incident in which they had been physically injured by their husbands following the first 6 months of treatment, and second, official police arrest records were used to determine the date of the first arrest for spouse abuse for the same period of time. The time between the end of treatment and the first incident of physical injury as reported by victimized wives, or the time between the end of treatment and the first arrest for spouse abuse as found in official records, was considered the survival time. When both sets of data were analyzed using the Kaplan-Meier technique (Norris, 1993), differences in survivorship profiles across groups were not

significant. The log-rank (Mantel-Cox) test for group differences was 1.35 ($p = .72$) for victim reports and 3.48 ($p = .32$) for arrest records.

Military Setting

It is possible that the military setting in which the experiment was conducted could, in and of itself, explain the no-difference findings if men in the Navy referred to the FAC for assaulting their wives perceived that their Navy careers would be put at risk if they continued to abuse their wives. The effects of such a perception may have overwhelmed the deterrent effects of the interventions. This point of view is consistent with a "stake in conformity" interpretation of the findings (Toby, 1957).

Two sets of analyses were completed to assess this hypothesis. One assessment compared the frequency of continued spouse abuse for men in the control group who reported that their Navy careers were or were not important to them. The assessment found no statistically significant differences for wife reports of episodic abuse and just one significant difference when the MCTS was used. Likewise, no statistically significant differences were found for assessments that compared outcome data for men who reported that their Navy careers would or would not be damaged if their superiors were to learn that they were continuing to abuse their wives. These findings suggest that a referral to the FAC and the associated visibility to commanding officers may not have been the overwhelming deterrent to continued abuse hypothesized earlier.

Prior Abuse

When the data on offending for the 6-month period prior to referral to the FAC (which did not include the presenting incident) were compared with the same data for the 1-year period following the first 6 months of treatment (see Tables 1 and 2), significant reductions in both the prevalence and frequency of violence were found for men assigned to all four of the experimental groups. These reductions were recorded irrespective of the treatment group to which men were randomized and irrespective of whether husbands or wives reported the violence. Although not noted in Table 3, the MCTS subscale scores for the 6-month period prior to the referral to the FAC paralleled those reported for the episodic measures.

These findings may be explained, in part, by an occurrence known as "telescoping." With this phenomenon, which is often associated with retrospective measures, old events are recalled as occurring more recently than they actually occurred, thus artificially inflating the baseline scores. The findings could also be a consequence of conducting four consecutive interviews over an 18-month period. If the interviews associated with the research are conceptualized as treatment interventions, they could explain the reductions in violence noted for men irrespective of the experimental treatment group to which they were assigned. That is, men were required, as a result of the interviews, to think periodically about violence-related issues in a way in which they may not have thought if the interviews had not been conducted. The interviews may have become, in this manner, interventions. Whatever the explanation, these reductions do not appear to be the results of the treatment interventions of the experiment.

Discussion

Findings from this study indicate that the cognitive-behavioral model, as implemented, demonstrated little power to foster change in men receiving treatment for spouse abuse. All of the assessments made, including comparisons of victim reports of continued abuse, perpetrator reports of continued abuse, official arrest records, and survival analyses, point to the same conclusion: The interventions of the cognitive-behavioral model failed to produce meaningful changes in the behavior they were designed to impact.

Although the effects of a military setting may play some role in deterring Navy personnel who abuse their wives, the Navy cannot assume that a referral to the FAC represents a treatment in and of itself or that interventions administered by the Navy will not add to the deterrent effect of such a referral.

Further, it would not be unreasonable to argue that a military setting is an optimal place in which to evaluate the efficacy of batterer treatment programs. All of the men were required to attend treatment, most men with serious mental health problems had been screened out of the population, all the participants were literate and reasonably competent, all were married and gainfully employed, and alcoholism and drug addiction were closely monitored and addressed. These circumstances may lend credence to the findings of this experiment and foster consideration among practitioners and others from nonmilitary settings.

Finally, the results of this experiment underscore the need to use experimental designs to evaluate all interventions for domestic violence. As noted at the outset, although the results of dozens of evaluations of batterer treatment programs indicate that the majority of men referred for treatment for spouse abuse discontinue their physical abuse of their wives or cohabitant partners, almost none of the studies used experimental research designs to evaluate what appear to be treatment successes. This limitation seriously compromises the validity of their conclusions. When the wife-reported data of the present experiment, for example, are examined independent of the control group, the findings indicate that 83% of the men in treatment (men's, conjoint, and rigorous monitoring) did not reinjure their wives during a 1-year outcome period. This finding ranks among the best of the evaluations found in the literature. However, when these results are examined in the context of the behavior of the men assigned to the control group, it is apparent that the treatment interventions were not responsible for the relatively low recidivism rates. The use of strong research designs is imperative if interventions for men who batter are to ever be validated as effective.

Perhaps the greatest limitation of the San Diego Navy Experiment is the fact that it stands alone. The results of the five replications of the Minneapolis Experiment (Garner, Fagan, & Maxwell, 1995; Sherman, 1992) demonstrate, if nothing else, that the results from a single experiment cannot be generalized to the world. The results of the San Diego Navy Experiment could also be questioned because of the use of a one-size-fits-all approach to the treatment of men who batter. It is generally believed that men abuse their wives/partners for a variety of reasons and that interventions must be tailored to their differential motivations/needs. A one-size-fits-all approach to the treatment of batterers may not be expected to address the different treatment needs of all batterers.

The results of this experiment suggest at least three priorities for future research. First, replication is needed. Second, the call

throughout the past decade for the use of rigorous experimental designs to assess treatment interventions for men who batter should be taken seriously. The risks of conducting randomized experiments to assess interventions for men who batter are likely to be fewer than the consequences of failing to do so. Finally, the possibility that a one-size-fits-all approach to the treatment of men who abuse their wives/partners is responsible for the ineffectiveness of treatment should receive full and preferential attention. Offender types should be identified and matched to intervention approaches to be followed by rigorous experimental evaluation.

References

- Boruch, R. F. (1994). The future of controlled randomization experiments: A briefing. *Evaluation Practices, 15*, 265-274.
- Chalk, R., & King, A. (1998). *Violence in families: Assessing prevention and treatment programs*. Washington, DC: National Academy Press.
- Crowell, N. A., & Burgess, A. W. (1996). *Understanding violence against women*. Washington, DC: National Academy Press.
- Davis, R. C., & Taylor, B. G. (1997). A proactive response to family violence: The results of a randomized experiment. *Criminology, 2*, 307-333.
- Edleson, J. L., & Tolman, R. M. (1992). *Intervention for men who batter: An ecological approach*. Newbury Park, CA: Sage.
- Eisikovits, Z. C., & Edleson, J. L. (1989). Intervening with men who batter: A critical review of the literature. *Social Service Review, 63*, 385-414.
- Fagan, J. (1996). *The criminalization of domestic violence: Promises and limits* (NIJ Research Report). Washington, DC: U.S. Department of Justice, National Institute of Justice.
- Feder, L. (1998, July). *A test of the efficacy of court mandated counseling for domestic violence offenders: Preliminary results from the Broward County experiment*. Paper presented at the international conference Program Evaluation and Family Violence Research, Durham, NH.
- Garner, J., Fagan, J., & Maxwell, C. (1995). Published findings from the Spouse Assault Replication Program: A critical review. *Journal of Quantitative Criminology, 11*, 1, 3-28.
- Geffner, R. (with Mantooth, C.). (2000). *Ending spouse/partner abuse: A psychoeducational approach for individuals and couples*. New York: Springer.
- Hamberger, L. K., & Hastings, J. E. (1993). Court-mandated treatment of men who assault their partners. In N. Z. Hilton (Ed.), *Legal responses to wife assault* (pp. 188-229). Newbury Park, CA: Sage.
- Healey, K., Smith, C., & O'Sullivan, C. (1998). *Batterer intervention: Program approaches and criminal justice strategies*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, National Institute of Justice.
- Mollerstrom, W. W., Patchner, M. A., & Milner, J. S. (1992). Family violence in the Air Force: A look at offenders and the role of the Family Advocacy Program. *Military Medicine, 157*, 371-374.
- Neidig, P. H. (1985). *Domestic Conflict Containment Program workbook*. Beaufort, SC: Behavioral Science Associates.
- Neidig, P. H. (1986). The development and evaluation of a spouse abuse treatment program in a military setting. *Evaluation and Program Planning, 9*, 275-280.
- Norusis, M. (1993). *SPSS for Windows: Advanced statistics* (Release 6.0) [Computer software manual]. Chicago: SPSS.
- Palmer, S. E., Brown, R. A., & Barrera, M. E. (1992). Group treatment program for abusive husbands: Long-term evaluation. *American Journal of Orthopsychiatry, 62*, 276-283.
- Rosenfeld, B. D. (1992). Court-ordered treatment of spouse abuse. *Clinical Psychology Review, 12*, 205-226.
- Saunders, D. G. (1996). Feminist-cognitive-behavioral and process-psychodynamic treatments for men who batter: Interaction of abuser traits and treatment models. *Violence and Victims, 11*, 393-414.

- Saunders, D. G. (1999). Feminist, cognitive, and behavioral group interventions for men who batter: An overview of rationale and methods. In D. Wexler (Ed.), *Domestic violence 2000: An integrated skills program for men* (pp. 21–31). New York: Norton.
- Sherman, L. W. (1992). *Policing domestic violence: Experiments and dilemmas*. New York: Free Press.
- Straus, M. A. (1979). Measuring intrafamily conflict and violence: The Conflict Tactics (CT) scales. *Journal of Marriage and the Family*, 41, 75–88.
- Szinovacz, M. E. (1983). Using couple data as a methodological tool: The case of marital violence. *Journal of Marriage and the Family*, 45, 633–644.
- Toby, J. (1957). Social disorganization and stake in conformity: Complementary factors in the predatory behavior of hoodlums. *Journal of Criminal Law, Criminology and Police Science*, 48, 12–17.
- Tolman, R. M., & Bennett, L. W. (1990). A review of quantitative research on men who batter. *Journal of Interpersonal Violence*, 5, 87–118.
- U.S. Navy (1993a). *Domestic violence conjoint groups—A combined treatment program: Group leader's manual*. San Diego, CA: San Diego Naval Base, Family Advocacy Center.
- U.S. Navy (1993b). *Domestic violence men's group: Group leader's manual*. San Diego, CA: San Diego Naval Base, Family Advocacy Center.
- Weinstein, G. S., & Levin, B. (1989). Effect of crossover on the statistical power of randomized studies. *The Society of Thoracic Surgeons*, 48, 490–495.
- Wexler, D. (Ed.). (1999). *Domestic violence 2000: An integrated skills program for men*. New York: Norton.

Received January 25, 1999

Revision received October 5, 1999

Accepted October 8, 1999 ■



AMERICAN PSYCHOLOGICAL ASSOCIATION SUBSCRIPTION CLAIMS INFORMATION

Today's Date: _____

We provide this form to assist members, institutions, and nonmember individuals with any subscription problems. With the appropriate information we can begin a resolution. If you use the services of an agent, please do **NOT** duplicate claims through them and directly to us. **PLEASE PRINT CLEARLY AND IN INK IF POSSIBLE.**

PRINT FULL NAME OR KEY NAME OF INSTITUTION _____

MEMBER OR CUSTOMER NUMBER (MAY BE FOUND ON ANY PAST ISSUE LABEL) _____

ADDRESS _____

DATE YOUR ORDER WAS MAILED (OR PHONED) _____

CITY _____

STATE/COUNTRY _____

ZIP _____

☐ PREPAID ☐ CHECK ☐ CHARGE
CHECK/CARD CLEARED DATE: _____

YOUR NAME AND PHONE NUMBER _____

(If possible, send a copy, front and back, of your cancelled check to help us in our research of your claim.)

ISSUES: ☐ MISSING ☐ DAMAGED

TITLE _____

VOLUME OR YEAR _____

NUMBER OR MONTH _____

Thank you. Once a claim is received and resolved, delivery of replacement issues routinely takes 4–6 weeks.

(TO BE FILLED OUT BY APA STAFF)

DATE RECEIVED: _____

DATE OF ACTION: _____

ACTION TAKEN: _____

INV. NO. & DATE: _____

STAFF NAME: _____

LABEL NO. & DATE: _____

Send this form to APA Subscription Claims, 750 First Street, NE, Washington, DC 20002-4242

PLEASE DO NOT REMOVE. A PHOTOCOPY MAY BE USED.

Intervention Programs for Perpetrators of Intimate Partner Violence: Conclusions from a Clinical Research Perspective

CHRISTOPHER I. ECKHARDT,
PhD^a

CHRISTOPHER MURPHY, PhD^b

DANIELLE BLACK, BA^b

LAURA SUHR, BA^a

SYNOPSIS

In this article, the authors consider the empirical status of batterer intervention programs (BIPs) for male perpetrators of intimate partner violence (IPV). Recent reviews have reported only small average effect sizes for BIPs, with the small number of randomized trials showing little benefit of BIP attendance in preventing future abuse. The most widely adopted BIP intervention model has little empirical justification to support this dominance, yet states with standards governing the content of BIPs often mandate this approach as a contingency for state funding. Little data exist concerning the moderators and mediators of BIP effects on IPV recidivism, and a variety of factors threaten to impede future design advancements, including "turf" battles regarding the causes of IPV and limited funding outlets. Given this discouraging summary, the authors argue that research efforts concerning BIP effectiveness should borrow the design strategies and programmatic research efforts that have proven successful in psychotherapy research, in which significant advances have been made with regard to the evaluation and validation of empirically supported treatments for a wide variety of mental health problems. They conclude by calling for a new generation of IPV researchers to work across professional boundaries in a multidisciplinary manner to design the sophisticated evaluation studies that funding agencies would readily support, and that would provide the substantive answers to the many IPV-related public health questions that remain.

^aDepartment of Psychological Sciences, Purdue University, West Lafayette, IN

^bDepartment of Psychology, University of Maryland, Baltimore County, Baltimore, MD

Address correspondence to: Christopher I. Eckhardt, Department of Psychological Sciences, 703 3rd St., Purdue University, West Lafayette, IN 47907; tel. 765-494-6983; e-mail <eckhardt@purdue.edu>.

©2006 Association of Schools of Public Health

Battered women are at higher risk for depression, suicide, post-traumatic stress disorder, alcohol/drug abuse or dependence, and poor physical health relative to women in non-abusive relationships.^{1,2} The financial costs associated with intimate partner violence in the U.S. exceed \$5.8 billion each year.³ Given the staggering individual, interpersonal, and societal costs of intimate partner violence (IPV), it is critical to examine how this problem is being addressed in terms of intervention programs that target the abusive individuals who are responsible for these negative outcomes. In this article, we discuss recent empirical findings regarding the efficacy and effectiveness of batterer intervention programs (BIPs). During the course of this review, we will also attempt to answer several critically important questions: Do such programs actually reduce the likelihood of subsequent acts of IPV? Are particular methods of BIP intervention more effective than others? If these programs are indeed effective, how do they accomplish their effects? In addition, we will contrast what we know about these programs with what we've learned from a parallel area of research and inquiry: the effectiveness and efficacy of psychotherapy for mental health problems. We'll compare and contrast the design and analysis strategies common to both approaches and outline possible strategies for improving research on BIP effectiveness.

CONCLUSIONS FROM PREVIOUS REVIEWS OF BIP EFFECTIVENESS

Research over the previous 20 years concerning the effectiveness of batterer intervention programs suggests that batterer intervention programs result in a small average reduction in intimate partner violence. An early review of 25 studies conducted by Rosenfeld to investigate the effectiveness of BIPs did not find promising results for the effect of treatment.⁴ Rosenfeld examined studies that reported recidivism rates and found that men who were arrested but not referred for treatment had a recidivism rate (39%) that was not significantly higher than the recidivism rate for men who were arrested and received treatment (36%). Davis and Taylor calculated the average effect size for five studies using quasi- or true experimental designs and obtained a small-to-moderate treatment effect size ($h=0.41$).⁵ A larger review of 17 studies conducted by Levesque and Gelles (1998) also reported small effect sizes for BIP (range: $h=0.18-0.27$).⁶

Two recent meta-analytic reviews of BIP effectiveness confirm that BIP attendance is associated with small effects on abuse recidivism. Babcock, Green, and Robie reviewed 22 studies that used quasi- or true experimen-

tal designs and police or partner reports of violence recidivism and found that the effects of BIP on violence cessation were small, with effect sizes ranging from $d=0.09$ to $d=0.34$.⁷ Using the more conservative effect size estimate, Babcock et al. concluded that men mandated to attend batterer intervention programs are only *5% less likely to commit an act of violence* against partners than men who do not attend/receive BIP.⁷ In addition, there appears to be an inverse relationship between research design complexity and effect size. Specifically, using partner reports of violence recidivism, the effect sizes associated with the influence of BIP on violence cessation based on studies using randomized experimental designs ($d=0.09$) were significantly lower than effect sizes from studies using non-randomized (quasi-experimental) designs ($d=0.34$). Using police reports of violence, randomized experiments resulted in smaller effects ($d=0.12$) than quasi-experiments ($d=0.23$), although these effects sizes were not significantly different from one another.⁷

Feder and Wilson also reviewed BIP effectiveness studies in a meta-analysis, including the experimental studies quantitatively reviewed by Babcock et al., but including only those quasi-experimental studies that established initial equivalence between groups via matching or use of statistical controls.⁸ Feder and Wilson's analysis of the 10 studies that qualified for inclusion indicated that among experimental studies, BIP had no overall effect on victim reports of physical violence ($d=0.01$) and small/moderate effects on official reports of spousal assault ($d=0.26$). Among quasi-experimental studies, BIP had a small iatrogenic effect according to victim reports of violence ($d=-0.11$) and official reports relative to no treatment comparison groups ($d=-0.14$). Interestingly, when BIP dropouts were used as a comparison condition in quasi-experiments using official reports of spousal assault, BIP had a very strong effect on violence reduction ($d=0.97$). While these meta-analytic reviews may be criticized for the relatively small number of studies reviewed, the general pattern of findings suggest that (1) BIP may be associated only with small positive effects in reducing the likelihood of male-to-female IPV, and (2) effect sizes attributable to BIP completion appear to decrease as a function of the methodological and evaluative rigor used in BIP effectiveness research.

Given these findings, it may be instructive to examine findings from randomized controlled trials (RCTs) of BIP effectiveness more closely. Taylor, Davis, and Maxwell randomly assigned 376 men convicted of a domestic violence offense in Brooklyn, NY, to an eight-week traditional gender-role focused group BIP (the "Duluth Model"), a 26-week traditional group BIP, or

a control (community service) group.⁹ Follow-ups at six-months and 12-months post-sentencing showed a significant impact of treatment on violence cessation according to police reports of violence recidivism. On partner reports of violence, however, there was no significant impact of treatment. In addition, when the treatment groups were separated by length of treatment, only men assigned to the 26-week treatment group had significantly fewer incidents of violence than those in the control group.⁹ Feder and Dugan randomly assigned 404 men in Broward County, FL to a Duluth Model BIP group plus probation monitoring, or to a probation monitoring-only control group.¹⁰ At a six-month and 12-month follow-up, the men assigned to receive BIP had levels of violence similar to the control group, according to police and partner reports. Dunford randomly assigned 861 men in the U.S. Navy stationed in San Diego to either a 26-week cognitive-behavioral group BIP, a 26-week couples therapy group, a rigorous monitoring group, or to a no-treatment control group.¹¹ Follow-up reports from female partners of male participants gathered six and 12 months post-treatment indicated no differences in male-to-female physical aggression across the four groups. Similarly, in a study of over 300 IPV perpetrators randomly assigned to treatment as a pre-trial diversion, treatment as a condition of probation, or a purely legal intervention (i.e., fine or jail time), Ford and Regoli found no significant difference in partner reports of violence across the three groups at six-month follow-up.¹² In contrast, Palmer, Brown, and Barrera found a positive effect of treatment with an abusive sample.¹³ Palmer et al. used a block random procedure to assign 59 Canadian IPV perpetrators to a 10-week psychoeducational treatment group versus a no-treatment control group. Police reports at a follow-up one year later indicated that men assigned to the treatment group showed lower rates of IPV recidivism than men in the control group.¹³ Overall, the results of these experimental investigations, which incorporate the highest degree of control over confounding factors, suggest that the BIP interventions produce, at best, quite modest benefits relative to non-BIP comparison conditions.

It is also important to examine whether there is evidence concerning differential treatment effectiveness: Does one type or format of BIP outperform other formats? At the risk of overgeneralizing this issue, there are two major models of treatment upon which most BIPs are based. The most popular intervention is a psychoeducational model that attempts,¹⁴ in a group format, to educate men about their attitudes concerning their perceived right to use power and violent coercion to

control or subjugate women (the "Duluth Model").¹⁵ A second approach, the cognitive-behavioral treatment (CBT) model, considers IPV as a learned behavior and focuses on the therapeutic modification of faulty cognitions and intense emotions and in teaching communication skills and emotion control techniques to prevent future violent behavior.¹⁶ The Babcock et al. meta-analysis found no significant difference between effect sizes of different BIP approaches.⁷ However, as noted by Babcock et al. and Healey et al., it has become increasingly difficult to discern between BIP groups that label themselves a Duluth Model program versus those that are self-described Cognitive-Behavioral Programs, a situation that artificially increases the likelihood of null results when comparing differences between intervention types.^{7,14} As both interventions purport to target faulty attitudes and beliefs and to address the behavioral consequences of holding those beliefs, it is common for both types of BIP formats to label themselves with the more generic CBT label. However, as noted by Murphy and colleagues, the primary differences between these formats lie in the methods used to enact cognitive and behavioral change.^{16,17} Duluth Model-based programs are typically didactic and education/consciousness-raising groups that consistently focus on issues relating to gender egalitarianism and patriarchal ideology. While there is a focus on attitudes and behaviors that is within the general scope of CBT practice, these programs typically have limited, if any, focus on coping with intense emotions, relationship skill building, trauma recovery, or other interventions to address various individual psychological problems. CBT interventions, in contrast, are more likely to address these latter topics, and also focus on the modification of faulty cognitive processes from a perspective based on the specific set of therapeutic principles and practices derived from generic cognitive behavioral treatments for mental health problems,¹⁸ rather than a strict focus on gender-themed factors (See Murphy and Eckhardt,¹⁹ for additional discussion of these approaches).

Aside from Duluth Model and CBT approaches, there are precious few alternative BIP formats. Despite the consistent link between disturbances in anger experience and expression among IPV perpetrators (for a review, see Norlander and Eckhardt²⁰), interventions that focus on anger control are not standard approaches and are often discouraged in states with guidelines governing BIP content.^{14,21} Interventions that focus on relationship systems are also forbidden according to the majority of state batterer intervention standards, despite extensive data demonstrating the bidirectional and often mutual nature of physical

aggression in close relationships.^{22–27} More research is needed using a treatment comparison design strategy that attempts to uncover whether there are meaningful differences in IPV-related outcomes across different BIP programming formats. The available studies have provided valuable insights in need of further empirical study.^{28–30} As an example, consider the empirical status of couples' treatment for partner abuse. In Dunford's randomized study of BIP (described above), there was no difference in IPV recidivism between men assigned to couples treatment versus a CBT group.¹¹ Using couples volunteering for treatment at a university marital distress clinic, O'Leary et al. found no difference between men assigned to either couples treatment versus a group Duluth Model intervention.²² Brannen and Rubin reported similar results using a court-referred sample.³¹ Thus, one can either conclude that couples treatment is unwarranted since it does no better than more traditional group treatments, or one can perhaps see couples treatment as a useful alternative for some violent couples (especially those who are clearly planning on staying together), since it appears to work just as well as traditional interventions.

A final complicating factor concerns the outcome variables measured in previous studies of BIP effectiveness, which have traditionally focused on dichotomous outcomes such as the cessation of physical violence. While partner-directed physical abuse has about a 20%–30% prevalence rate,^{2,32} it has a relatively low frequency of occurrence within violent couples.² As such, additional abuse indicators are necessary in order to capture more frequent but similarly damaging forms of IPV. One such indicator that has unfortunately received only infrequent attention in BIP evaluation research is psychological abuse. As indicated by a variety of studies, psychological abuse has a higher base rate of occurrence than physical abuse and is strongly associated with a variety of negative outcomes among victims.^{33,34} While a number of recent studies examining BIP effectiveness have used multiple measures of partner abuse (including psychological aggression) and relatively long follow-up assessment period of 1–2 years,^{11,29} these advances are relatively recent design developments; the majority of studies to date have utilized limited measures of IPV-related outcomes and relatively short follow-up assessment periods (i.e., six months post-intervention). Among these latter studies, it is possible that one reason for the lack of significant differences between treatment conditions is the low base rate of partner physical abuse and the lack of attention to other forms of IPV. This is not to say that psychological abuse should replace physical abuse as an outcome variable, but regular inclusion

of emotional abuse measures may add more useful information about higher base rate forms of IPV in addition to more difficult to predict and lower base rate physical forms of IPV that may go undetected in BIP evaluations with limited follow-up assessments.

BIPs were designed to go beyond mere incarceration or legal punishment in cases of IPV and provide an intervention that might actually change perpetrators' behavior for the long term and prevent future abuse from occurring. The available data concerning the effectiveness of such programs in actually accomplishing this goal are rather discouraging: a large percentage of men (around 40%–60%) either do not attend or drop out of BIP, and there is only a negligible relationship between attending BIP and IPV cessation.³⁵ It is our position that while data regarding BIP effectiveness have improved in many significant ways in recent years, much is simply unknown about how such programs should be designed, how they should be applied in the field, and how they should be studied empirically. This state of affairs is very similar to the status of psychotherapy for mental health problems approximately 40 years ago. At that time, there were a small number of treatments, with each claiming success and theoretical dominance. There were remarkably few carefully designed studies of treatment packages that used random assignment to treatment, construct valid assessment instruments, and long-term follow-up practices. Since that time, however, a tremendous amount has been learned concerning the effectiveness of psychotherapy, in large measure due to the development of rich and complex methodologies for evaluating the causal effects and generalizability of therapeutic interventions. The literature on psychotherapy and behavior change has much to offer the nascent area of IPV intervention and prevention programs in terms of theory, research design, and application.

DETERMINING BIP EFFECTIVENESS: THE VIEW FROM PSYCHOTHERAPY RESEARCH

Clinical research on psychotherapy and behavior change can be divided into two general categories: outcome research and process research. Outcome research focuses on the ability of an intervention to produce a targeted change in behavior, emotion, or symptoms. Process research focuses on the mechanisms of change or the ingredients of treatment that are associated with positive outcomes. Outcome studies have been further subdivided into efficacy research, which focuses on intervention effects in highly controlled experiments, versus effectiveness research, which focuses on outcomes in naturalistic, real-world

practice, and often has fewer controls in place to rule out alternative explanations of findings.³⁶

The last quarter-century has witnessed a steady stream of methodological advances in psychotherapy research, such that the modern clinical trial has become a very ambitious and expensive undertaking.³⁷ Long gone are the days when a small number of loosely diagnosed cases could be subjected to a novel intervention with clinician ratings of change from baseline as the primary outcome. To be considered state-of-the-art, the modern randomized controlled trial (RCT) requires, at minimum, the following features:

1. A sufficient number of cases to detect modest intervention effects (high statistical power).
2. Careful screening and diagnosis to ensure that study participants have the clinical problem under investigation and do not have other problems or complicating factors likely to compromise treatment efficacy (which, together with #1, implies access to a large potential subject population).
3. One or more well-specified experimental treatments with a manual (typically book-length), containing detailed conceptual and procedural information on competent treatment delivery.
4. Methods for demonstrating the integrity of treatment, at a minimum containing measures of adherence to treatment protocols and often also containing measures of competence (e.g., ratings of therapist skill in delivering the specified treatment) and discriminability (e.g., demonstration that specific elements of treatment are delivered exclusively or primarily within specific experimental conditions or that specific elements are not delivered within certain conditions).
5. Methods for training therapists and certifying their competent delivery of experimental treatments (often involving detailed training protocols and close supervision of training cases).
6. Specification of one or more control conditions that will promote causal conclusions regarding treatment efficacy, such as waiting list or no-treatment controls, placebo controls that receive a "theoretically inert" form of clinical attention, or treatment-as-usual controls.
7. Random assignment to experimental versus control conditions (often including additional procedures to promote balancing of key subject features across conditions).
8. Multiple reliable and valid measures of key outcome and process variables, preferably from more than one reporter (e.g., self-report, partner report, criminal justice data, clinician ratings) and generated from multiple methods (e.g., behavioral observation of treatment, self-monitoring) and gathered at pre-intervention baseline and multiple time points during follow-up in order to detect initial change, change trajectories, and maintenance of gains.
9. Data gatherers (assessors) who are not involved in the delivery of treatments and ideally are blind to treatment condition (to reduce the likelihood that participants will respond to please treatment providers or that assessors will bias data gathering to confirm hypotheses).
10. Detailed tracking strategies and incentives to reduce subject attrition in order to avoid erroneous experimental findings as a result of differential drop-out from conditions.
11. Sophisticated data processing and analytic techniques to handle missing data, identify individual growth/change functions, and assess the statistical significance, magnitude, and clinical significance of experimental effects.

Needless to say, such an undertaking is not for the faint-hearted, (and probably not for the untenured). To be considered "empirically supported," a treatment must have solid evidence of efficacy from at least one randomized controlled trial meeting most of the features listed above.³⁸ One can also set a higher standard, considering a treatment to be "empirically valid," which implies replicated success in RCTs conducted by more than one investigator at multiple sites, thus supporting strong claims regarding both efficacy and generalization of treatment effects.

To date, *there are no interventions for partner violence perpetrators that approach the standard of "empirically valid," and it is debatable whether any intervention can be labeled "empirically supported."* The majority of existing studies lack random assignment to treatment versus control conditions, and therefore cannot rule out alternative explanations such as spontaneous (naturally occurring) change in behavior over time (e.g., in pre-post-intervention designs with no control or comparison group) or selection artifacts (i.e., preexisting differences between treatment and controls that explain results).³⁹ Selection effects are prominent alternative explanations for findings that use drop-outs or treatment refusers as the control group, as these individuals are likely to differ from treatment completers in a number of important ways (e.g., lower stake in conformity, more disorganized life styles, greater antisocial features) that may account

for poorer outcomes. As noted above, the meta-analytic review by Babcock et al. indicated that the vast majority of studies used some type of non-randomized research design, mostly involving treatment drop-out controls.⁷ Among the five studies that used randomized experimental designs, very small average intervention effects were observed.

In response to the outcomes of experimental investigations of BIP effectiveness, various authors have highlighted limitations present among these studies, with some researchers suggesting that these flaws may be responsible for the otherwise negligible effects reported. Criticisms have been discussed either in terms of specific design flaws within a particular study or in terms of supposed limitations of RCTs in general,^{40–42} and the interested reader may wish to consult these sources for a more complete accounting of the issues involved. For example, while the research by Dunford applied most of the ideals of the RCT listed above, one may question the exclusive use of military personnel and the high rate of duty-reassignment among men in the study as it relates to the ability of the interventions to produce differential outcomes and the generalizability of findings to civilian populations.¹¹ More generally, Gondolf has questioned whether RCTs should indeed be the “gold standard” for determining research design quality, noting the practical limitations to implementing such designs in criminal justice settings and conceptual issues concerning how best to analyze differences among experimental and comparison groups (e.g., intent to treat versus completer analyses).⁴² Thus, acknowledging and addressing limitations to available research is a critically important step toward improving the quality of future BIP effectiveness research. However, it would be mistaken logic to imply that since experiments are imperfect and difficult to implement, less rigorous research designs are somehow more desirable. In addition, most of the research design elements offered as alternatives to randomized experiments (e.g., statistical modeling of treatment exposure in instrumental variables analysis) were developed to better approximate experimental designs with non-experimental data, and can be applied to strengthen the findings from experiments with problems such as subject attrition. Thus, it is a far stretch of reasoning to argue that these strategies are somehow superior to the RCT when it comes to generalized causal inference.³⁹

That said, it is worth highlighting the strong points present in the small number of BIP effectiveness studies that have approximated the ideals of the RCT, or that have at least used strong quasi-experimental

controls. Researchers have randomly assigned very large samples (e.g., 300–800) of men to competing intervention conditions,^{5,11,43} used multiple measures and reports from multiple informants (for a review see Gondolf⁴²), implemented assessments of treatment integrity and adherence,^{11,30} incorporated important analytical innovations from the public health arena including instrumental variable and propensity score analyses,^{41,44} and utilized long-term follow-ups of one to three years.^{11,41,30} Thus, in the previous decade, high quality experimental and quasi-experimental studies of BIP effectiveness have successfully informed the field about what may work in the context of programs designed to stop men’s abuse of their female partners. However, recent research syntheses cast some doubt on the overall degree of effectiveness among BIPs compared to relevant comparison conditions. Even if one were to accept the most optimistic appraisal of the effectiveness of partner violence interventions,^{42,45} serious questions remain regarding alternative explanations of intervention effects as a function of research methodology, and there is no clear way to attribute any potential effects to specific aspects of intervention. In a nutshell, both non-randomized studies and experimental studies with high attrition rates and inadequate treatment specification, which together account for the vast majority of research on BIP interventions, require that a number of plausible alternative explanations be ruled out before any causal conclusions of treatment efficacy can be drawn.

THE PROCESSES OF CHANGE IN BIP: THE VIEW FROM PSYCHOTHERAPY RESEARCH

A number of factors have emerged quite consistently as successful predictors of treatment outcome in the general research literature on psychotherapy and behavior change. Perhaps most notable is the working alliance—the collaborative relationship between therapist and client.⁴⁶ The alliance is typically thought to have three components: a warm bond between therapist and client, agreement on the goals of treatment, and agreement on the tasks or strategies needed to attain those goals. The collaborative alliance is a controversial concept in partner violence intervention, as supportive and empathic therapist behaviors thought to promote a strong alliance have been seen by some in the field as promoting collusion with the abuser’s negative outlook. It is also important to note that the alliance, as currently conceived, is not a therapist-delivered entity, but rather a relationship level, two-person construct. Thus, it is expected that both the client and therapist

contribute to the establishment of the working alliance, and that client personal characteristics may impede alliance formation.

To date, the available empirical evidence consistently supports the predictive value of a strong working alliance in partner violence intervention. In both a study of couples and gender-specific group treatments for voluntary (self-referred) abusive men as well as a study of cognitive-behavioral group treatment for primarily court-referred abusive men,^{47,48} ratings of the working alliance predicted lower levels of self-reported and partner-reported abusive behavior post-treatment. Interestingly, in the study by Taft and colleagues,⁴⁸ therapist ratings of the alliance late in group treatment (at sessions 11 and 13 of a 16-session program) were the most strongly associated with outcome (as compared to client ratings and early session ratings of both client and therapist). With this often interpersonally challenged and treatment-resistant population, it may take a while for the alliance to develop or for the therapist to have a clear picture of the quality of the alliance.

Additional factors that have been found to predict successful change in other areas of psychosocial treatment research have also enjoyed some predictive success in partner violence treatment. Compliance with homework assignments in CBT (i.e., participation in active change strategies) was associated with lower levels of psychological abuse after treatment. In addition, client ratings of positive group cohesion were associated with lower levels of both physical and psychological abuse at follow-up.⁴⁸

Client motivational readiness to change may play an important part in these process results. Researchers investigating the transtheoretical (“stages of change”) model of behavior change have reported that approximately one-third of abusive men mandated to attend BIP present with characteristics suggestive of the earliest stage of the behavior change process, i.e., men who do not recognize the existence of a problem and who have no plans to make active attempts at behavior change (the “precontemplative” stage).^{49–52} Precontemplative men reported using fewer behavior change processes than men in other stages,⁵⁰ reported fewer benefits relative to costs of making a commitment to nonviolence,⁵¹ demonstrated minimal therapeutic change over the course of BIP,⁵² and were more likely to be arrested for any criminal offense one year post-adjudication.⁵³ Not surprisingly, therefore, motivation to change is a strong predictor of the working alliance.⁵⁴ Abusive clients who reported higher motivational readiness to change at program intake established a stronger working alliance, which in turn is associated

with higher compliance with the structured change elements of treatment and with lower levels of post-treatment abusive behavior.¹⁹

Consistent with the process findings, the use of motivational interviewing (MI) strategies during the intake process at a community agency for partner violence treatment appears to promote client engagement into treatment and improvement on process factors related to outcome. Motivational interviewing was designed for working with substance abusing clients who are often ambivalent about change.^{55,56} It uses a high level of reflective listening, affirmation of client autonomy and control over the change process, techniques for “rolling with resistance,” and interventions tailored to the client’s stage of change.^{51,52} A study comparing a standard structured intake to an intake process involving two 45-minute motivational interviews found that abusive clients who received the motivational intake articulated more positive statements about treatment and took more personal responsibility for their abusive actions during early sessions of their subsequent domestic violence group (as measured by observational coding of group treatment sessions), had much higher levels of CBT homework compliance, and higher therapist ratings of the working alliance in subsequent group treatment.⁵⁷

In brief, the available research to date, although limited, indicates that partner-violent clients are quite similar to other psychosocial treatment populations in responding to therapist support and reflective empathy. Factors that predict successful outcomes in other areas of psychotherapy and behavior change likewise appear to predict cessation of physical assault and reduction of emotionally abusive behavior in this treatment population. Although we do not as yet have sound empirical support for this speculation, careful reading of many existing treatment manuals in this area indicate that high levels of therapist confrontation and critical or punitive attitudes toward abusive clients by service providers may impede the development of the working collaborative alliance and other active elements of the helping relationship.^{58–60}

DESIGN STRATEGIES AND ASSUMPTIONS IN IPV RESEARCH

Research on the effectiveness of interventions for IPV perpetrators has generally followed the logic of social policy analysis (trying to investigate the effectiveness of existing programs in the field) rather than the logic of clinical trials (trying to develop interventions and test their efficacy in carefully controlled settings before

attempting to generalize them to field work). While the social policy approach may offer important answers to short-term needs existing within the criminal justice and advocacy communities to develop useful violence rehabilitation and desistance practices, it offers relatively little for our long-term understanding of how violence cessation programs ought to be designed and implemented, and how these practices actually work to affect behavior change. Thus, state-of-the-art technology from clinical behavior change research has not been sufficiently brought to bear on counseling/treatment interventions for partner-violent men.

But why not? Is it any more difficult to successfully study what interventions promote nonviolent change than to investigate what treatments alleviate episodes of depression or reduce the likelihood of self-injurious behavior? While there are indeed a variety of practical obstacles that make randomized designs more difficult to implement in criminal justice settings (see below), such obstacles are not insurmountable; public health researchers have clearly demonstrated that close links between basic and applied research can result in successful prevention programs for a wide variety of health perils. Thus, there are other forces at work besides task difficulty that likely explain why research on BIP effectiveness is lagging behind. We offer several possibilities:

1. The research/practice rift

There has been, and continues to be, a deep and often contentious controversy surrounding how to conceptualize the causes of, and intervention approaches for, IPV. A perspective held by many practitioners and grass roots activists suggests that the root causes of IPV lay in our fundamentally patriarchal societal and institutional structures that tacitly or overtly reward the continued domination of males over females, and that justify any means (including physical aggression), enabling men to occupy positions of power.⁶¹ Males absorb these messages of male privilege during socialization by community and family members, and apply them in intimate relationships in the form of behaviors that exert power, control, and domination over their female partners. Given this heuristic, which represents a starting point in the development of structured intervention programs for IPV perpetrators, it follows that intervention strategies should be centered around psychoeducational reprogramming, whereby the patriarchal ideologies and philosophy of male privilege among perpetrators are exposed, power and control tactics discouraged, and more gender-egalitarian strategies encouraged. Most existing intervention programs and state coalitions against domestic violence espouse

both this specific conceptual framework and singular intervention approach. In some states, approaches that run counter to these perspectives are discouraged or prohibited by existing standards.^{14,24}

A somewhat different perspective often espoused by IPV researchers and empirically-oriented practitioners from a wide array of professional backgrounds is that power, control, and misogynistic attitudes are indeed important factors in understanding IPV and intervening with perpetrators. However, additional factors are added to the equation that may also be important in understanding the causes of IPV and in designing intervention strategies. Thus, this perspective provides allowances for such perpetrator-focused factors as psychopathology, anger arousal disturbances, cognitive distortions, and the long-term effects of childhood traumas, and considers alternative intervention strategies for abusive men that focus on additional risk factors or treatment modalities (e.g., conjoint treatment) other than those specified by the patriarchal ideology model. These intervention targets are usually supported by available empirical studies, but are often criticized by practitioners for running too far afield from the feminist analyses of IPV causation.²⁴

While it seems obvious that the perspectives of practitioners and researchers are intertwined and largely complementary, the more proximal consequences of this rift have unfortunately been twofold: (1) a pervasive lack of trust between BIP practitioners/administrators and IPV researchers, and (as a consequence), (2) an infrequent collaboration between IPV researchers and BIP practitioners. It would seem logical for BIP practitioners/administrators to adopt the most efficacious intervention for batterers available, since doing so is directly related to the one goal that unifies all parties involved: promoting the future safety and welfare of abused partners. Researchers ought to be in a prime position to aid BIP programs in this regard, as the question of what works best for whom is the kind of question that researchers, who are more likely to have a specific background in research design and methodology, are ideally suited to answer. And since researchers have more direct access to external funding sources, they could financially aid BIP programs such that those agencies could perhaps serve more perpetrators more effectively. But researchers are unlikely to progress with research questions relating to improvements in BIP effectiveness without the aid of BIP practitioners/administrators, who typically have the greatest access to IPV perpetrators and thus serve as quasi-gatekeepers in the path of BIP research efforts. While reports have suggested the potential benefits of researcher-practitioner collaborations in BIP design

and evaluation,²⁹ cooperation between researchers and practitioners at present needs strengthening.⁴²

2. Difficulties with definition and design

As noted above, most BIP research follows the logic of social policy analysis: evaluating the criminal justice policies currently in practice in response to IPV and the effects that these policies have on future acts of violence. While such an approach is warranted given the emerging nature of this research area, we would posit that a critical question still unaddressed by the social policy approach is a question of definition: Is a BIP punishment, a therapeutic intervention, an educational experience, or some amalgam of all three? Clearly, IPV offenses are violations of the law and the criminal justice system reacts to that offense by mandating that most offenders attend a batterers intervention group or face a jail penalty, so in that sense BIP is obviously a punishment in response to a specific criminal offense. But we expect individuals in these groups to learn something, to gain important information that helps them understand their prior decisions to act violently and contributes to a new understanding and personal responsibility to eliminate IPV from their lives. So, in that sense, BIP is psychoeducational. Yet, it's also clear that a variety of psychological and lifestyle factors are among the more robust risk factors for acts of IPV, and many BIP programs will address these risk factors in a therapeutic manner (e.g., cognitive behavioral interventions that change faulty beliefs or improve emotional coping). BIP is therefore difficult to define as an amalgam of punishment, education, and therapy.

This amalgamation creates problems from a design and evaluation standpoint. It is difficult to disentangle which of these three operations are contributing to any changes that occur as a result of BIP in designing and evaluating a particular intervention. While any therapeutic intervention involves an amalgam of ingredients that promote change including client education, the presumed active components of the intervention, and other nonspecific factors stemming from the therapeutic relationship, BIP presents a most unique case given the coercive nature of the referral source, the often restricted nature of what events are allowed to serve as the causes of IPV as well as the focus of BIP interventions,²⁴ and the ever-present threat of legal retribution for treatment noncompliance. To disentangle these components, one must design the evaluation in such a way that educational, therapeutic, and punishment factors that may be mediating the relationship between BIP attendance and IPV outcomes can be reliably assessed. To date, there have been relatively few efforts to measure the mediators of BIP

outcomes, and doing so would add extremely valuable information about the processes of change involved in IPV cessation. Recent articles have provided a detailed description of how researchers can assess mediating and moderating factors in RCTs, and BIP effectiveness research must rise to this design challenge to address three pressing and unanswered research questions regarding BIP effectiveness: (1) How can particular interventions be tailored to address the most critical treatment targets for IPV perpetrators? (2) Under what conditions and for whom does a given intervention work? (3) Do the presumed active components of a particular BIP modality uniquely contribute to behavior change?⁶² We would argue that the research design strategies and lessons learned from decades of research on psychotherapy and behavior change hold substantial promise for answering these questions, but have received only minimal attention to date.

Since there is little agreement about how best to define BIP, it follows that there are diverse viewpoints regarding how best to study its effects. Most would agree that the RCT is the ideal method of determining the efficacy of a clinical intervention, but as noted above, there have only been a handful of such studies in the history of BIP research and evaluation, and they are the studies with some of the lowest effect sizes for reducing violent recidivism. There are several reasons for the dearth of RCTs evaluating BIPs. Given the long list of RCT-defining aspects listed above, one can readily infer that the complexity inherent to such designs involves numerous practical obstacles within any research context. In the criminal justice context, a variety of 'real world' demands and problems can interfere with implementation of an RCT. A variety of community, legal, and criminal justice agencies are typically involved or implicated with the design and implementation of the research, and each will likely have a broad array of concerns, differing perspectives, and other interests that must be addressed before the study can proceed. For example, will BIP program administrators and/or domestic violence court judges allow random assignment of batterers to treatment and control groups? Will there be an opportunity to present informed consent information to potential participants at the beginning of a BIP session, or must this occur at another time or location? Will there be sufficient cooperation between the researchers and the agencies that control access to important information (e.g., police, probation officers, BIP counselors) to allow the research design to be fully implemented? In addition, the dynamic changes of personnel, policies, and procedures that typically occur in criminal justice settings over the course of an evaluation period can significantly

affect how design details are implemented.⁴³ Such practical demands and organizational constraints require a close degree of cooperation between researchers and nonresearchers well in advance of data collection or even submission of a grant application. Without clear communication and detailed cooperative agreements between the parties involved, there will be a limited degree of trust between the research and nonresearch personnel, which in turn may lead to difficulties related to how the research is conducted and ultimately lead to the presence of factors that may confound the results of the evaluation.

3. Funding sources

Even with a strong working relationship between a BIP agency and one or more clinical researchers, an RCT comparing, for example, just two variations of BIP to a suitable control group requires considerable resources.³⁷ Costs include intervention staff, assessment staff, money for participant and partner follow-up, and funds to support scientific administration and project oversight. Assuming the design is well constructed and the overall proposal is conceptually clear, it *should* be the case that such a strong research coalition investigating a critically important public health problem will have little difficulty locating a funding source. However, this is far from true. In fact, the limited funding available for BIP-related research threatens to attenuate the already limited progress of this area of research. Since IPV perpetration is not an official diagnosis in the latest revision of the Diagnostic and Statistical Manual of Mental Disorders (i.e., DSM IV-TR),⁶³ it may fall outside the realm of traditional behavior change funding sources (e.g., the National Institute of Mental Health). While various branches of the National Institutes of Health (NIH) have funded research on IPV, a recent electronic search of IPV-related grants funded by NIH since 2000 indicate that less than 20% of funded projects actually investigated specific intervention programs to reduce IPV. In addition, it has become increasingly the case that new research proposals need to substantively address factors relating to mental health (i.e., using DSM-IV diagnostic categories) in addition to violent behavior to be favorably reviewed. Other branches of the U.S. Health and Human Services, most notably the Centers for Disease Control and Prevention, have recently become more regular and visible funding outlets for IPV research and BIP investigations in particular.

The creation of the Violence Against Women Office (VAWO) within the United States Department of Justice in 1995 was designed to provide a central coordinating agency that would enact and enforce the monitoring

and criminal justice policies relating to the 1993 Violence Against Women Act. In addition, in collaboration with the research arm of the Department of Justice, the National Institute of Justice (NIJ), it has also served as a specialized funding source for research on IPV, with the number of discretionary grants awarded to IPV-related research project increased from 92 funded proposals in 1995 to more than 400 in 2000.⁶⁴ NIJ/VAWO has funded numerous important research projects regarding BIP, published a series of useful monographs concerning the present and future status of BIP,^{14,40,65} and convened several noteworthy national conferences devoted solely to discussing the state of BIP research and practice. However, there has been limited interest expressed in recent requests for proposals in investigating specific intervention strategies for IPV perpetrators. While NIJ/VAWO has been interested in the effectiveness of BIPs and has funded two of the five published studies using random assignment of perpetrators to treatment, the focus of such studies has been on the criminal justice or social policy implications of BIP rather than on the specific content of the interventions involved. In addition, the amount and duration of NIJ grants are often well below that required to conduct a high quality RCT investigation.

Thus, in order to secure funding, a research coalition may need to go beyond the rudimentary questions of whether BIP works and differential effectiveness of multiple BIP interventions, and instead make their BIP evaluation proposal relevant to issues relating to mental health/substance abuse or directly relevant to criminal justice policy and practice. The problem, of course, is that accumulated research on BIP effectiveness and our knowledge of the potential moderators and mediators of its effects are at such an elementary stage that these very basic questions are precisely what need to be answered at the present time. Thus, as researchers gather more evidence about the processes of change underlying IPV cessation and put this information to use in RCTs investigating BIP process and outcome, there is also the coexisting dimension that such research needs to be designed in the context of additional content areas deemed important by relevant funding agencies.

CONCLUSIONS

In conclusion, the limited research on BIP effectiveness and the lack of suitable application of sophisticated research design strategies that have so clearly benefited research on psychotherapy and behavior change are not because of a lack of awareness that these issues exist; rather, any careful examination of the general

BIP literature suggests that it is an area where theoretical/ideological concerns have largely outstripped the importance of empirical evidence. For example, some have argued that state standards governing BIP content appear to have been formulated largely on the basis of loyalty to a particular explanatory model rather than on a careful examination of the research evidence on abuse perpetrators or evidence for a particular intervention model's empirical support.^{25,66} While there is likely to be more flexibility in program design at the local level,⁴² additional empirically informed insights are needed to guide the optimal development of BIP interventions.

As a result of these problems, research on interventions for IPV perpetrators is at a critical juncture. The accumulation of largely unremarkable outcomes regarding BIPs could potentially signal to those in the criminal justice community that such programs are simply not worth the effort: Why mandate an intervention that men have little motivation to attend, that at best has a small impact on criminal recidivism, and that doesn't really qualify as a punishment, an educational experience, or a therapeutic intervention? Should the criminal justice community decide to eliminate BIPs, funding sources would shrink even further and research would stagnate. Thus, we hope that the relatively discouraging findings from controlled studies to date do not completely undermine our efforts to develop effective interventions for abusers.

Perhaps a more optimistic implication is that the field needs an infusion of new researchers, with parallel increases in support from relevant funding agencies. Thus, if a concerted effort could be made to cast aside issues of professional boundaries and theoretical ownership of the area, and to apply the most sophisticated methods for the study of intervention and behavior change with support from multiple funding agencies, the field might be in a position to do something that it cannot do today: to answer the questions posed at the outset of this paper—do such programs actually reduce the likelihood of subsequent acts of IPV? Are particular methods of BIP intervention more effective than others? What are the moderators and mediators of BIP success and failure; how do BIPs accomplish their effects? Perhaps a new generation of clinical researchers can bring us to a point where we can actually possess empirically informed answers to these questions from theoretically supported and methodologically sophisticated research designs. These are the kinds of answers that ought to inform matters of criminal justice policy, rather than the very limited and much debated conclusions that are currently in play. To get there, funding agencies will have to increase their funding of this kind

of high level research in order to generate not only a new generation of researchers but a new generation of more complex research questions.

For example, one promising avenue is to examine the effects of batterer interventions for co-occurring conditions that often accompany partner violence. While some elements of the BIP community continue to reject the notion that individual mental health problems can ever influence IPV, it is reasonable to assume that at least a small percentage of perpetrators will also have co-occurring psychiatric and substance-related problems. While these problems may or may not directly cause IPV, they certainly make the intervention process much more complicated. Indeed, researchers are just starting to progress on elucidating the prevalence and role of comorbid mental health problems and partner violence,⁶⁷ and this should presumably involve examination of effects that interventions for those particular disorders have on partner violence. In fact, Gondolf and colleagues at Indiana University of Pennsylvania are currently conducting an NIJ-funded randomized trial of adjunctive mental health interventions for IPV perpetrators, the results of which are eagerly anticipated.

In conclusion, it is clear at this point that researchers investigating intervention programs for men who abuse intimate partners can no longer “go it alone”—the multifaceted nature of IPV and the complex issues surrounding the treatment goals of BIPs mandate nothing less than a multidisciplinary approach to addressing this problem. While this is hardly a novel suggestion, putting this ideal into practice has proven challenging. In our view, the noteworthy methods and lessons learned from the area of psychotherapy research has much to offer the emerging area of BIP effectiveness research. Individuals who have been charged with the responsibility of counseling men mandated to BIP, be they at treatment delivery or administrative levels, must move beyond the traditional and well accepted perspectives on how such programs should be designed. While this is always a difficult goal to achieve, the evidence reviewed in this article suggests that current BIP programs adopted by most jurisdictions are in need of improvement. Thus, it is time for researchers and practitioners to move beyond the traditional boundaries that have separated various professional areas involved in IPV research and treatment in hopes that future generations of individuals can experience love and intimacy without abuse.

REFERENCES

1. Golding JM. Intimate partner violence as a risk factor for mental disorders: a meta-analysis. *J Fam Violence* 1999;14:99-132.

2. Straus MA, Gelles RJ. How violent are American families? In: Straus, MA and Gelles, RJ, editors. *Physical violence in American families: risk factors and adaptations to violence in 8,145 families*. New Brunswick (NJ): Transaction; 1990.
3. Centers for Disease Control and Prevention (US), National Center for Injury Prevention and Control. *Costs of intimate partner violence against women in the United States*. Atlanta: CDC; 2003.
4. Rosenfeld BD. Court-ordered treatment of spouse abuse. *Clin Psychol Rev* 1992;12:205-26.
5. Davis RC, Taylor BG. Does batterer treatment reduce violence? A synthesis of the literature. *Women and Crim Justice* 1999;10:69-93.
6. Levesque DA, Gelles RJ. Does treatment reduce recidivism in men who batter? A meta-analytic evaluation of treatment outcome. Paper presented at the International Family Violence Research Conference, Durham, NH; 1998 Jul.
7. Babcock JC, Green CE, Robie C. Does batterers' treatment work? A meta-analytic review of domestic violence treatment. *Clin Psychol Rev* 2004;23:1023-53.
8. Feder L, Wilson DB. A meta-analytic review of court-mandated batterer interventions programs: can courts affect abusers' behavior? *J Exp Criminol* 2005;1:239-62.
9. Taylor BG, Davis RC, Maxwell CD. The effects of a group batterer treatment program: a randomized experiment in Brooklyn. *Justice Q* 2001;18:171-201.
10. Feder L, Dugan L. A test of the efficacy of court mandated counseling for domestic violence offenders: the Broward experiment. *Justice Q* 2002;19:343-75.
11. Dunford FW. The San Diego Navy experiment: an assessment of interventions for men who assault their wives. *J Consult Clin Psychol* 2000;68:468-76.
12. Ford DA, Regoli MJ. The criminal prosecution of wife batterers: process, problems, and effects. In: Hilton NZ, editor. *Legal responses to wife assault*. Newbury Park (CA): Sage Publication; 1993. p. 127-64.
13. Palmer SE, Brown RA, Barrera ME. Group treatment program for abusive husbands: long-term evaluation. *Am J Orthopsychiatry* 1992;62:276-83.
14. Healey K, Smith C, O'Sullivan C. *Batterer intervention: program approaches and criminal justice strategies*. Washington: National Institute of Justice; 1998.
15. Pence, E, Paymar, M. *Education groups for men who batter: the Duluth Model*. New York: Springer; 1993.
16. Daniels JW, Murphy CM. Stages and processes of change in batterers' treatment. *Cogn Behav Pract* 1997;4:123-45.
17. Murphy CM, Baxter VA. Motivating batterers to change in the treatment context. *J Interpers Violence* 1997;12:607-19.
18. Dobson, K, editor. *Handbook of cognitive-and behavioral therapies*, 2nd ed. New York: Guilford Press; 2001.
19. Murphy CM, Eckhardt CI. *Treating the abusive partner: an individualized cognitive-behavioral approach*. New York: Guilford Press; 2005.
20. Norlander B, Eckhardt CI. Anger, hostility, and male perpetrators of intimate partner violence: a meta-analytic review. *Clin Psychol Rev* 2005;25:119-52.
21. Maiuro RD, Hagar TS, Lin H, Olson N. Are current state standards for domestic violence perpetrator treatment adequately informed by research? A question of questions. *J Aggress Maltreatment Trauma* 2002;5:21-44.
22. O'Leary KD, Heyman RE, Neidig PH. Treatment of wife abuse: a comparison of gender-specific and conjoint approaches. *Behav Ther* 1999;30:475-506.
23. Archer J. Sex differences in aggression between heterosexual partners: a meta-analytic review. *Psychol Bull* 2000;126:651-80.
24. Dutton DG, Nicholls TL. The gender paradigm in domestic violence research and theory: Part 1—The conflict of theory and data. *Aggress Violent Behav* 2005;10:680-714.
25. Murphy CM, O'Farrell TJ. Couple communication patterns of maritally aggressive and nonaggressive male alcoholics. *J Studies on Alcohol* 1997;58:83-90.
26. O'Leary KD, Barling J, Arias I, Rosenbaum A, Malone J, Tyree A. Prevalence and stability of physical aggression between spouses: a longitudinal analysis. *J Consult Clin Psychol* 1989;57:263-8.
27. Stets JE, Straus MA. Gender differences in reporting marital violence and its medical and psychological consequences. In: Straus MA, Gelles RJ, editors. *Physical violence in American families: risk factors and adaptations to violence in 8,145 families*. New Brunswick (NJ): Transaction; 1990. p. 151-65.
28. Edleson J, Syers M. The relative effectiveness of group treatments for men who batter. *Soc Work Res Abstr* 1990;26:10-7.
29. Gondolf E, Yllo K, Campbell J. Collaboration between researchers and advocates. In: Kaufman G, Jasinski J, editors. *Out of darkness: contemporary research perspectives on family violence*. Thousand Oaks (CA): Sage Publications; 1997. p. 255-61.
30. Saunders DG. Feminist-cognitive-behavioral and process-psychodynamic treatments for men who batter: interaction of abuser traits and treatment models. *Violence Vict* 1996;11:393-414.
31. Brannen SJ, Rubin A. Comparing the effectiveness of gender specific and couples groups in court-mandated spouse abuse treatment programs. *Res Soc Work Pract* 1996;16:252-60.
32. Schafer J, Caetano R, Clark CL. Rates of intimate partner violence in the United States. *Am J Public Health* 1998;88:1702-4.
33. Murphy CM, Hoover SA. Measuring emotional abuse in dating relationships as a multifactorial construct. *Violence Vict* 1999;14:39-53.
34. O'Leary KD. Psychological abuse: a variable deserving critical attention in domestic violence. *Violence Vict* 1999;14:3-23.
35. Daly JE, Pelowski S. Predictors of dropout among men who batter: a review of studies with implications for research and practice. *Violence Vict* 2000;15:137-60.
36. Seligman ME. The effectiveness of psychotherapy: the Consumer Reports study. *Am Psychol* 1995;50:965-74.
37. Kazdin AE, editor. *Methodological issues and strategies in clinical research*. Washington: American Psychological Association; 2003.
38. Chambless DL, Hollon SD. Defining empirically supported therapies. *J Consult Clin Psychol* 1998;66:7-18.
39. Shadish WR, Cook TD, Campbell DT. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin; 2002.
40. Jackson S, Feder L, Forde DR, Davis RC, Maxwell CD, Taylor BG. *Batterer intervention programs: where do we go from here?* Washington: Dept. of Justice (US), Office of Justice Programs, National Institute of Justice; 2003.
41. Gondolf E, Jones A. The program effect of batterer programs in three cities. *Violence Vict* 2001;16:693-704.
42. Gondolf EW. Evaluating batterer counseling programs: a difficult task showing some effects and implications. *Aggress Violent Behav* 2004;9:605-31.
43. Gondolf EW. *Culturally-focused batterer counseling for African-American men* [cited 2006 Jan 9]. Final report to the National Institute for Justice; 2005. Available from: URL: http://www.ojp.usdoj.gov/nij/vawprog/vaw_portfolio.pdf
44. Jones A, D'Agostino R, Gondolf E, Heckert A. Assessing the effect of batterer program completion on reassault using propensity scores. *J Interpers Violence*. In press 2006.
45. Gondolf EW. *Batterer intervention systems: issues, outcomes, and recommendations*. Thousand Oaks (CA): Sage Publications; 2002.
46. Bordin ES. The generalizability of the psychoanalytic concept of the working alliance. *Psychother Theory Res Pract* 1979;16:252-60.
47. Brown PD, O'Leary KD. Therapeutic alliance: predicting continuance and success in group treatment for spouse abuse. *J Consult Clin Psychol* 2000;68:340-5.
48. Taft CT, Murphy CM, King DW, Musser PH, DeDeyn JM. Process and treatment adherence factors in group cognitive-behavioral therapy for partner violent men. *J Consult Clin Psychol* 2003;71:812-20.
49. Prochaska JO, DiClemente CC, Norcross JC. In search of how people change: applications to addictive behaviors. *Am Psychol* 1992;47:1102-14.
50. Eckhardt CI, Babcock J, Homack S. Partner assaultive men and the stages and processes of change. *J Fam Violence* 2004;19:81-93.
51. Levesque DA, Gelles RJ, Velicer WF. Development and validation of a stages of change measure for men in batterer treatment. *Cognit Ther Res* 2000;24:175-200.
52. Scott KL, Wolfe DA. Readiness to change as a predictor of outcome in batterer treatment. *J Consult Clin Psychol* 2003;71:879-89.
53. Eckhardt CI, Holtzworth-Munroe A, Norlander B, Sibley A, Cahill M. Readiness to change, partner violence subtypes, and treatment outcomes among men in treatment for partner assault. *Violence Vict*. In press 2006.
54. Taft CT, Murphy CM, Musser PH, Remington NA. Personality,

- interpersonal, and motivational predictors of the working alliance in group cognitive-behavioral therapy for partner violent men. *J Consult Clin Psychol* 2004;72:349-54.
55. Miller WR, Rollnick S. Motivational interviewing: preparing people to change addictive behavior. New York: Guilford; 1991.
 56. Miller WR, Rollnick S. Motivational interviewing: preparing people to change addictive behavior. 2nd ed. New York: Guilford; 2002.
 57. Musser PH, Semiatin JN, Taft CT, Murphy CM. Motivational interviewing as a pre-group intervention for partner violent men. Manuscript under review for publication; 2005.
 58. Murphy CM, Baxter VA. Motivating batterers to change in the treatment context. *J Interpers Violence* 1997;12:607-19.
 59. Scalia J. Psychoanalytic insights and the prevention of pseudosuccess in the cognitive-behavioral treatment of batterers. *J Interpers Violence* 1994;9:548-55.
 60. Taft CT, Murphy CM. The working alliance in intervention for partner violence perpetrators: recent research and theory. *J Fam Violence*. In press, 2006.
 61. Dobash RE, Dobash R. Violence against wives: a case against the patriarchy. New York: Free Press; 1979.
 62. Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. *Arch Gen Psychiatry* 2002;59:877-83.
 63. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-IV-TR. 4th ed. Washington: American Psychiatric Association; 2000.
 64. Government Accounting Office (US). Problems with grant monitoring and concerns about evaluation studies. Washington: GAO; 2002.
 65. Saunders DG, Hamill RM. Violence against women: synthesis of research on offender interventions. Washington: National Institute of Justice; 2003.
 66. Dutton DG. Transforming a flawed policy: a call to revive psychology and science in domestic violence research and practice. *Aggress Violent Behav*. In press 2006.
 67. Kessler RC, Molnar BE, Feurer ID, Appelbaum M. Patterns of mental health predictors of domestic violence in the United States: results from the National Comorbidity Survey. *Int J Law Psychiatry* 2001; 24:487-508.

RELATIVE EFFECTIVENESS OF GROUP TREATMENTS FOR MEN WHO BATTER

Edelson, Jeffrey L., Syers, Maryann, *Social Work Research & Abstracts*, 01480847, Jun90, Vol. 26, Issue 2

ABSTRACT:

An experiment was conducted to compare six different group treatment programs for men who batter. The 283 men included in the study were randomly assigned to one of three forms of group treatment offered in two different intensities. Of these men, 153 completed 80 percent or more of their assigned programs. Six months after group treatment ended, 92 program completers or their partners were located and interviewed. Analyses of the resulting data revealed that shorter, more structured group treatment was most effective in reducing the number of men reported violent and using terroristic threats during follow-up. The majority of men in all six programs were reported to be using less severe threats during follow-up.

Group treatment for men who batter has been available in the United States since the late 1970s. Several national surveys have found the number of such programs to have grown dramatically in recent years (Feazell, Mayers, & Deschner, 1984; Roberts, 1982). As programs have proliferated, so have the treatment methods advocated by both activists and practitioners.

Programs vary in organizational structure, ties to the criminal justice system and the battered women's movement, and the group format by which treatment is delivered. Most treatment programs vary along major dimensions of group process and both the total number and the intensity of group sessions. The degree to which one program is seen as more effective than another often generates heated debate and hinges on practitioners' beliefs about the group process and number of sessions necessary to achieve desired outcomes.

Debates about the effectiveness of one or another group process often focus upon the degree of structure group leaders provide. Group structure appears to vary along a continuum from highly structured, educational programs (Pence & Paymar, 1986; Reilly & Grusznski, 1984) to minimally structured, self-help groups (Goffman, 1980; Jennings, 1987). Most programs in North America offer group treatment lying somewhere

between these two endpoints (Rosenbaum, 1986; Saunders, 1984; Tolman & Edleson, 1989).

The recommended number of group sessions also is a source of debate. Some practitioners advocate "long-term" treatment of 1 to 5 years (Ewing, Lindsey, & Pomerantz, 1984). However, most batterers' treatment groups in North America appear to be "time limited." In their survey of programs in the United States, Feazel et al. (1984) found that batterers' treatment programs averaged 6 weeks in length. Other practitioners have advocated slightly longer programs. For example, Sonkin, Martin, and Walker (1985) presented a 12-week program, whereas Brygger and Edleson (1987) described a 16-week program that met for 32 sessions.

The debate about the number and intensity of services offered is affected by financial pressures within the insurance industry that have led to the imposition of time limits on treatment reimbursed through health plan coverage. A persistent shortage of funding for domestic violence services places additional pressure on men's treatment programs to be as efficient as possible.

In the decade since the first batterers' groups were offered, limited research has been reported that sheds light directly on issues of optimal group structure and intensity of service. Only in recent years have systematic evaluations of group treatment for batterers become available. Generally, they have shown group treatment to be effective in ending violence among 59 to 84 percent of program completers over short follow-up periods and in achieving desired changes on measures of anger, depression, attitudes towards women, jealousy, and communication skills. The same studies generally have found that a majority of program completers have used threats of violence against their partners during follow-up (Eisikovits & Edleson, 1989).

These evaluations are grounds for some optimism, but they shed little light on the debates about the relative effectiveness of varying group structures and number of services offered. Several authors (Gondolf, 1987; Rosenbaum, 1988; Saunders, 1988) have highlighted the need for new research to compare these various treatment modalities.

Activists in the battered women's movement (Hart, 1988) also are calling for greater accountability and monitoring of batterers' programs. Although these calls are not aimed directly at generating new studies, research that closely monitors treatment

programs and ends in the dissemination of outcome data may greatly increase information about the effects of treatment on batterers' behavior.

The research reported here was designed to compare common group treatment modalities and to offer practitioners more detailed answers to questions concerning optimal group structures and numbers of sessions. These general questions were translated into two specific ones that guided the design of this study: (1) What are the relative effects of three group treatment models upon batterers' postgroup use of violence and threats of violence? and (2) What are the relative effects of two different intensities of time-limited treatment upon batterers' postgroup use of violence and threats of violence? The study also examined how the selected models of treatment interact with selected intensities to create outcomes unique to specific model-intensity combinations.

In the research presented here, batterers were assigned randomly to receive one of six possible treatment conditions. Victim and perpetrator interviews were completed 6 months after the end of group treatment. The results of these follow-up interviews are the focus of this article.

METHOD

Subjects

A total of 283 men were included in the study sample. The following descriptive data are based on participants' self-reports of demographic characteristics. The total number of men reporting varies slightly by variable. The men ranged in age from 18 to 57 years with a mean age of 31.8 years ($SD = 8.1$). Of those reporting racial or ethnic status, the majority of men (73.7 percent, $n = 193$) were white and the remainder were men of color; 30 (10.6 percent) were black, 10 (3.8 percent) were native American, 7 (2.7 percent) were Hispanic, and 1 (0.4 percent) was Asian American.

The level of education attained by the men in the sample ranged from 4 to 12 years. The mean number of years of education completed was 12.7 ($SD = 2.2$). Half of the men (50.2 percent, $n = 126$) were employed full time at intake, whereas a large percentage (33.5 percent, $n = 84$) were unemployed. The majority (53.5 percent, $n = 144$) reported an annual income below \$10,000, and only 8.2 ($n = 22$) percent reported an annual income above \$30,000.

Just over one-third (34.5 percent, $n = 96$) of the men reported themselves married at intake, compared with one-fourth (24.5 percent, $n = 68$) who were separated and one-fifth (20.1 percent, $n = 56$) who were single and had never been married. Almost 4 out of 10 men (39.1 percent, $n = 108$) reported having been violent with previous partners, whereas a similar proportion (42.8 percent, $n = 113$) reported having been violent with strangers.

A minority of the men (38.3 percent, $n = 102$) were ordered to treatment by the courts, whereas the rest entered treatment voluntarily, but usually under some type of social pressure -- for example, a partner had entered a shelter for battered women, received an order for protection, or filed for divorce. About half (50.7 percent, $n = 137$) of the men reported having previously received some chemical dependency treatment. A similar proportion (50.2 percent, $n = 127$) of the men reported receiving mental health treatment sometime before intake.

Most men reported being fathers (87.5 percent, $n = 238$) with the average number of children being 2.2 ($SD = 1.8$). About one in five (21.6 percent, $n = 57$) reported having been violent toward their children. Most (68.6 percent, $n = 190$) had witnessed violence in their families as children, and about the same proportion (68.2 percent, $n = 189$) had been abused physically as children.

Treatment Procedures

The three models of treatment compared in this study were (1) an education model, (2) a self-help model, and (3) a combined model that integrated education and self-help. Each of these models was offered in two intensities.

Treatment Models. The three models of group treatment reflected major trends in the North American service network. The education model relied heavily on lectures, videotaped and role-played demonstrations, and short group discussions. The professionally trained facilitators were called "teachers," and the group members were called "students." Five modules were presented over the length of the group: (1) "Introduction," (2) "Abuse: How It Happens," (3) "Abuse: Its Impact on People in My Life," (4) "Why Has Abuse Become Part of My Life?" and (5) "How To Change." A workbook with regular readings and assignments accompanied each module. Group discussions were kept to a minimum and focused directly upon the material being

presented rather than upon detailed analysis of personal events taking place in the men's lives.

Several sectors have been increasing pressure for the use of educational programs in batterers' treatment. The view that violent behavior by men is learned and socially reinforced has long been a tenet of the battered women's movement (Dobash & Dobash, 1979). Changing such behavior is viewed as requiring, in part, a reeducation of men who batter rather than a psychotherapeutic change in personality. Separately, criminal justice systems often refer those convicted of driving while intoxicated and other offenses to short, structured programs for education. Pressures for similar programs have increased as the criminal justice system has become more involved in the issue of woman battering. Perhaps the most widely disseminated educational model is a 24-session program developed by Pence and Paymar (1986). Reilly and Grusznski (1984) described a similar but shorter educational program.

At the other end of the continuum of treatment models, a minimally structured, self-help group was developed in which members defined topics covered and a former batterer facilitated meetings. The facilitator was required to have been nonviolent for at least 1 year and to have received extensive training in leading a self-help group. All facilitators used in this study previously had led similar groups. A trained professional who acted as a consultant also was present at each meeting, but only in a backup role to the facilitator. In each meeting men "checked in" and asked to "take time" to discuss personal events or issues. The group members defined the topics to be discussed; however, four topics were discussed by either the facilitator or the consultant at some point during the treatment: (1) personal responsibility for violent behavior, (2) developing a personal plan for being nonviolent, (3) use of "time out" as a way to diffuse tension, and (4) how violence develops along a cycle (Walker, 1979). This content was also presented in the other two models.

Self-help groups are widely available in North America, with perhaps the best-known groups being Alcoholics Anonymous and Parents Anonymous. These groups are member run and focus on supplying mutual support and sharing personal stories. Most Parents Anonymous groups have a professional sponsor present for consultation and advice. Goffman (1980) was one of the first to propose batterers' self-help groups and advocated a model similar to Parents Anonymous. More recently, Jennings (1987)

argued that unstructured groups may provide a superior environment for encouraging men to change.

The self-help model used here was designed to be similar to the large number of self-help programs being offered to batterers throughout North America. In most cases, self-help is applied as an aftercare program for men who first complete more structured programs. In this study, however, self-help was used as the primary intervention for the men randomly assigned to receive it.

In the middle of the continuum between the education and self-help models was the combined model. It offered men the opportunity both to receive educational lectures and to take time to discuss personal issues and events in detail with the group. Generally, men checked in at the beginning of the groups, and portions of most group meetings then were devoted to educational presentations, followed by personal time to tell stories and work on personal issues.

This model was intended to represent group programs most widely used across North America. In the literature describing groups for men who batter, there are many examples of groups combining educational and self-help procedures in single programs (Rosenbaum, 1986, Saunders 1984; Tolman & Edleson, 1989).

The education and combined groups included the introduction of a great deal of information. This content was kept very similar between the two programs, although the combined model presented the content in less detail, to allow time for group discussion of individual problems. As mentioned above, several basic pieces of information were introduced to all groups, including the self-help groups.

All three models are representative of a large number of programs in North America. The specifics of each program were designed and implemented by the staff of only one agency, the Domestic Abuse Project, but drew upon the experiences of other practitioners, programs, battered women, and activists.

In both the education and the self-help models, group membership was limited to those entering at fixed times, with new members being added only once during any man's tenure in the group. In the combined model, men entered at the start and remained together as a group until the end of the program without new members being added.

All group facilitators, teachers, and leaders were men. All had worked with batterers' groups before being involved in this study and were closely supervised to ensure similar application of material within each model tested. The two teachers in the education model also led some of the combined model groups. The reverse was not true: the professional consultants to the self-help group did not participate in leadership of other groups. Thus, group leader differences were not controlled.

An average of seven men participated in each group. In all three models men were offered the opportunity for one-to-one crisis intervention and referral to additional counseling as group leaders deemed necessary.

Intensities of Treatment. Each of the three treatment models was offered in two intensities: 12 and 32 sessions. The 12-session groups of each model met weekly for 2 hours and 15 minutes over 12 consecutive weeks. The 32-session groups of each model met twice weekly for 2 hours and 15 minutes over 16 consecutive weeks. The 12-session version involved 27 hours of group contact, compared with the 32-session version's 72 group contact hours. Total contact hours of the 12-session groups were 37.5 percent of those received by men in the 32-session groups.

All of the programs studied here fell under the general rubric of time-limited treatment. They differed primarily in intensity (once versus twice per week) and only slightly in length (12 versus 16 weeks). However, the differences in staff resources devoted to the 12-session and the 32-session programs were considerable.

Study Design

Six groups of each model-intensity combination were offered. Each month, three new groups were offered, one of each model. For the first 6 months of the study, each model was offered in a 32-session format, and for the second 6 months, each model was offered in a 12-session format. A total of 36 groups were conducted over the 12 months from June 1986 through May 1987. Every man who contacted the agency during this period was assigned randomly to one of the three treatment models being offered. During any single month, however, groups were offered in only one of the two intensities.

This design resolved many pragmatic concerns of agency staff regarding the conduct of a randomized field experiment within a complex agency. The ability to compare 12-

and 32-session programs and interactions between treatment models and intensities was weakened by offering one intensity first and another later. Although men had an equal opportunity to be randomly assigned to one type of treatment model, splitting the study into two intensities resulted in the men not having an equal chance to be assigned to 12- or 32-session versions of treatment. The date a man contacted the agency determined the intensity of the program he would enter. There is little likelihood, however, that certain months of the year attract a specific type of man more than other months. This study did not include a control group of men who did not receive services or who were placed on a waiting list.

Measurement Procedures

Data were collected on violence, threats of violence, and a number of other variables. These data were collected at intake, at closing, and 6 months after group treatment ended.

Before treatment, two intake questionnaires -- a demographic information sheet and an intake questionnaire -- were completed. The demographic information sheet was a form that requested information on each client's referral source, current relationship status, and demographic characteristics, such as education, race or ethnicity, religion, and income. The intake questionnaire was completed by the counselor during a pre-group intake interview with the man. It included an extensive set of questions on the man's violent and threatening behavior and a thorough history of his current and past behavior. The violence and threat questions asked during the intake interview were similar to those used in the Conflict Tactics Scale (Straus, 1979). At intake, men were required to respond to 29 questions about three categories of violence or threats of violence: (1) threats of violence (stomped out in the middle of an argument, screamed at or insulted his partner, interrupted her sleeping or eating, restricted her physical movement or social contact, verbally pressured his partner for sex, physically disciplined children, or threatened to leave his partner), (2) terroristic threats of violence (physically harmed pets; threatened to hit or throw something at his partner; threw, hit, or smashed objects; or drove recklessly to frighten his partner), and (3) physical violence (burned, pushed, grabbed or shoved his partner, slapped or spanked his partner with an open hand, bit or scratched his partner or hit her with something, physically forced his partner to do something, physically forced his partner to have sex, punched his partner with his fist, kicked his partner, punched partner in her stomach

when she was pregnant, threw his partner bodily, beat his partner unconscious, choked or strangled his partner, threatened or actually used a weapon against his partner). Each question was rated on a 9-point scale, with 0 being "never in the past 6 months" and 9 being "several times a day." The identical questions were asked of men and their partners during follow-up interviews.

When a man ended treatment at the agency, his group counselor completed a closing form that asked for information about violence and threats reported during the treatment period, childhood victimization reported during treatment, the types and amounts of services provided, and minimal information about the man's status at closing. In addition, the counselor was required to rate the man's prognosis for success on a series of variables.

Six-month follow-up interviews were conducted by graduate social work students. Priority was given to contacting the man's most recent intimate partner as listed in agency documents. Partner reported data were given priority because research indicates that women report violence more often than do their male partners (Edleson & Brygger, 1986; Jouriles & O'Leary, 1985; Szinovacz, 1983). Women's reports were considered a more conservative estimate of program effects than data collected from men. If a female partner was not available at follow-up, the interviewers contacted the man. The follow-up interviews involved a structured series of questions. Most important were a set of questions, identical to those at intake, that asked about the recurrence of violence, terroristic threats, and threats in the 6 months since the end of group treatment.

Other questions, regarding chemical use, relationship status, and satisfaction with agency services, were asked. Each interview was conducted by telephone and required approximately 20 minutes to complete.

RESULTS

The results presented in this section were derived primarily from 6-month follow-up interviews, with additional information gathered from intake and closing data. Analyses of differences among study subjects are presented, followed by analyses of treatment model and intensity effects.

Subject Differences

Two hundred eighty-three men were included in the final sample, of which 153 completed 80 percent or more of their assigned groups and 130 dropped out of the programs before completing at least 80 percent. One hundred thirty-three of the men or their partners (47 percent of the sample) were located and completed the 6-month follow-up telephone interviews. Ninety-five of the men or their partners found at follow-up were program completers, and 38 were noncompleters. Thus, 62 percent of the program completers were found at follow-up, whereas only 29.2 percent of the noncompleters were found.

The above statistics reflect a great deal of subject attrition. Such narrowing of the sample reduces the power of the study to provide generalizable findings by leaving open the possibility of a skewed follow-up sample. Such a large decrease in subjects between intake and follow-up is troubling but common for many treatment programs in North America. Many men who batter are not highly motivated to complete treatment. The ability at follow-up to find 62 percent of those who completed treatment or their partners compares favorably with other major studies in this field. For example, during a 6-month follow-up in the landmark police study of Sherman and Berk (1984), 49 percent of the sample were located.

The intake data for the 133 men (or their partners) who completed 6-month follow-up interviews were compared with the intake data for those who were not found. These two groups of men were not significantly different on most demographic variables. Those found at follow-up, however, tended to be better educated (chi-square = 5.17; df = 1, p = .02) and to be earning higher incomes (chi-squares = 6.31; df = 2, p = .04) at intake.

The differences found may represent differences between program completers and noncompleters rather than differences between those found and those not found at follow-up. Other research (Grusznski & Carrillo, 1988) has found that program completers tend to be better educated than those who drop out of treatment.

No differences existed between those found and those not found at follow-up on severity of violence as reported at intake, except on one variable. The men found at the 6-month follow-up were more likely to have reported themselves as having threatened to leave their partners (chi-square = 14.44; df = 1, p = .0001) before

intake than were the men not found. On all other questions related to threats of violence and actual violence at intake, the two groups were not significantly different.

Treatment Differences

The relative effects of different groups on postgroup use of violence, terroristic threats, and threats were central to this research project. Thus, the focus of the analyses is on comparing men in different programs who attended at least 80 percent of the sessions in the program to which they were assigned. Due to missing data on three completers interviewed at follow-up, data were available for 92 program completers. These data were provided by partners' in 80 cases and were self-reported by men in 12 cases. The cases in which only self-reported data were available were distributed evenly across all six possible model-intensity configurations of treatment.

The results of analyses of the 6-month follow-up data are presented below. First, two- and three-way cross-tabulations were conducted to examine the relationships between dependent variables (violence, terroristic threats, and threats) and independent variables (intensity of treatment, group model, race, marital status at intake and at follow-up, income, education and employment levels, whether or not the man was ordered by the court to seek treatment, extra services received while participating in group treatment, and prior mental health or chemical dependency counseling received). Near-significant relationships then were included in a logit analysis.

For final analyses, men were categorized into one of two groups on each of the three major dependent variables: (1) threatening or not threatening, (2) using terroristic threats or not, (3) and physically violent or not. The use of categorical data thus indicated the use of both chi-square and logit analyses. Logit analyses were used here in an effort to determine the combination of variables that best predicted the subjects' continued use of violence and terroristic threats.

Effects on Violence. Analysis of the relationship between intensity of treatment and postgroup reports of violence revealed that the 12-session groups, disregarding the type of treatment offered, resulted in approximately 10-percent lower rates of violence at follow-up than did 32-session groups (Table 1). These differences were not statistically significant, suggesting roughly equivalent effects of 12- and 32-session programs. Indeed, program completers of 12-session groups were less likely to be reported violent at follow-up than were those in 32-session ones.

A two-way cross-tabulation of treatment model by postgroup reports of violence indicated that, disregarding the number of group sessions offered, the more structured and educational a group, the less likely a man was to be reported violent at follow-up (Table 2). The structured education model seemed to have the greatest impact: only 32.3 percent ($n = 10$) of the men were reported violent at follow-up. The combined model, integrating education and self-help components, achieved similar results, with only 34.3 percent ($n = 12$) of the men reported violent at follow-up. The minimally structured self-help group had the least impact on violence, with more than half (53.8 percent, $n = 14$) of the participants reported violent at follow-up. These differences were not statistically significant.

Three-way cross-tabulations revealed statistically significant differences between treatment models offered in 12-session formats and their effects on violence reported at follow-up. Men participating in 12-session education and combined groups were much less likely to be reported violent at follow-up than were those who participated in 12-session self-help groups. Only 20 percent ($n = 3$) of those in the 12-session education groups and only 26.9 ($n = 7$) percent of those in the 12-session combined groups were reported violent at follow-up, compared with 64.3 percent ($n = 9$) of the men in the 12-session self-help groups who were reported violent at follow-up. These differences were statistically significant (chi-square = 7.55; $df = 2$, $p = .02$).

Minority men were less likely to respond favorably to 12-session groups than to 32-session groups. However, only 19 men of color were found at follow-up, 10 of whom participated in 12-session groups and 9 of whom participated in 32-session groups, making it impossible to draw conclusions based on these data. Of these few men, 70 percent ($n = 7$) of those who participated in 12-session groups and were found at follow-up were reported to be violent, compared with 26.2 percent ($n = 11$) of white men. This difference was statistically significant (chi-square = 5.05; $df = 1$, $p = .03$). Similar but non-significant trends were found in data from the 32-session groups. A total of 53.8 percent ($n = 14$) of the white men but only 22.2 percent ($n = 2$) of the men of color who participated in 32-session groups were reported to be violent at follow-up.

A logit analysis was performed, using ethnic or racial status, intensity of group, and group model as well as several other variables that achieved a liberal significance level of $p \leq .15$ in the cross-tabulations. In all, eight variables meeting this criterion were

included in the logit analysis: (1) ethnic status, (2) marital status at intake, (3) intensity of group, (4) group model, (5) previous mental health or chemical dependency counseling, (6) level of employment, (7) education, and (8) income.

When performing the logit analysis, none of these variables entered into the equation as predictors of continued, postgroup use of violence. The multivariate, logit analysis thus indicated that none of the variables listed above--including ethnic or racial status, group model, and intensity--was able reliably to predict continued violence at a statistically significant level. This finding casts doubt on the significant three-way interactions noted above.

Effects on Terroristic Threats. Differences existed among programs on reports of severe, terroristic threats (Table 3). Twelve-session groups, disregarding the type of group treatment, tended to show about a 6-percent greater impact on reducing the incidence of terroristic threats after treatment. In both the 12-session and the 32-session groups, however, a majority of men were reported to be continuing their use of terroristic threats during the 6 months after treatment. The differences between the intensities of treatment again were not statistically significant.

The final analysis compared the effects of different group treatment models on postgroup use of terroristic threats (Table 4). The differential effects of treatment models on terroristic threats were greater than those found when examining the violence variable. Men who participated in education groups were the least likely (38.7 percent, $n = 12$) to be reported to be using terroristic threats during follow-up. This rate of threatening behavior was lower than that found for the combined model (51.4 percent, $n = 18$, using terroristic threats) and lower than that reported for the self-help model (73.1 percent, $n = 19$, using terroristic threats). These differences were statistically significant ($\chi^2 = 6.79$, $p = .03$).

Three-way cross-tabulations did not reveal significant differences regarding interactions among treatment intensity, model, and terroristic threats. Sorting men by the intensity of the group in which they participated yielded the same trends as cross-tabulations by group treatment models. Within both the 12-session and the 32-session groups, the education model outperformed the others in reducing the likelihood of postgroup use of terroristic threats. When comparing each model in its 12-session version with the same model in its 32-session version, the less intensive model was

always more effective than the more intensive one in reducing the use of terroristic threats. The range of program effects was great. For example, only 33.3 percent ($n = 5$) of the men who participated in the 12-session education groups were reported to have used terroristic threats during the 6 months after group treatment, whereas 75 percent ($n = 9$) of the men in the 32-session self-help group were reported to have used terroristic threats during the follow-up period.

A cluster of three socioeconomic variables did reveal statistically significant interactions with terroristic threats and group model. In comparing group models, a man was significantly less likely to have been reported as using terroristic threats at follow-up if he participated in an education group and was full-time employed (chi-square = 7.13; $df = 2$, $p = .03$), had an income over \$10,000 (chi-square = 7.97, $p = .02$), and was educated beyond high school (chi-square = 14.16, $p = .0008$).

Cross-tabulations suggested nine potentially significant predictor variables, including the above socioeconomic ones, that were subsequently included in the logit analysis. These variables were similar to those used in the earlier logit and included group model, marital status at intake and at follow-up, prior mental health or chemical dependency counseling, whether or not the man was ordered by the court to have treatment, level of education, income and employment, and the amount of extra services received while participating in treatment. Again, the criterion for inclusion in the logit analysis was a cross-tabulation that yielded a chi-square with a liberal p -value of .15 or less.

Marital status at intake and model of group were the only variables that entered the logit equation (Table 5). Including these two predictor variables in the logit analysis yielded a model with a high degree of predictive ability, or goodness of fit (chi-square = 0.356, $p = .837$).

Only the group model and marital status variables remained in the equation. The positive coefficient for marital status at intake indicates that if the couple were not living together at intake, the man was less likely to be continuing his use of terroristic threats at follow-up. This seems reasonable, in that the opportunity to engage in threatening behavior is likely to be lower if the man and the woman are not living together. This inference is tempered somewhat by the finding that marital status at follow-up did not enter the equation here, and neither of these marital status variables

entered the equation on the violence variable. The coefficients for group model indicate that a man was less likely to be reported to be using terrorist threats during follow-up if he was in an education group as opposed to a combined group, a combined as opposed to a self-help group, or an education as opposed to a self-help group.

Effects on Threats. The same analyses were attempted for data gathered on postgroup use of other, less severe, threats. All but 14 men were reported to be continuing some type of threat, such as stomping out of an argument. These 14 men were distributed evenly across different model-intensity combinations. Because there were so few men reported not to be threatening during follow-up, more detailed analyses were impossible.

CONCLUSION

This study was designed to discover more about optimal treatment models and intensities for helping batterers end their violence and threats of violence. On average, 12-session treatment groups were as effective as 32-session groups in reducing repeat incidents of violence and terroristic threats as reported at follow-up. Men participating in education groups were significantly less likely to be reported to be continuing their use of terroristic threats during the 6-month follow-up period. Although there were similar trends regarding program effects on reports of violence, these differences were not significant.

The conclusion that brief, time-limited treatment can be as effective as more intense treatment is consistent with research on other social service populations (Butcher & Koss, 1978; Johnson & Gelso, 1980; Reid & Shyne, 1969; Videka-Sherman, 1988). The reduction of treatment intensity offers the potential for more efficient use of treatment dollars and staff resources. One explanation for this finding may be that men in 12-session groups accelerated their efforts to make the greatest use of the shorter amount of time available. Also, a greater percentage of men in the 12-session groups attended 80 percent or more of group sessions. Possibly, more men in the 32-session groups dropped out before completion after judging themselves successful.

The more structured education model most consistently appeared to reduce postgroup reports of both violence and terroristic threats. More structured educational programming offers the possibility of a more easily transferred treatment technology.

Training new practitioners to use educational programs that rely on packaged lectures, demonstrations, videotapes, and participant handbooks may require fewer hours of instruction and result in more consistent program implementation.

Responses to the combined model seemed mixed. Men were less likely to be reported violent, but more likely to be using terroristic threats during the follow-up period. On average, men who participated in self-help groups were much less successful in ending their violence and use of terroristic threats during follow-up. However, this study did not examine the use of self-help groups in aftercare. Such groups most often are offered as maintenance programs after more structured group treatments end. The results of this study argue against the use of self-help groups as the primary form of treatment, but they offer no indication of their usefulness after a man's participation in a more structured program.

Previous outcome studies of group treatment for men who batter have found from 16 to 46 percent of men located at follow-up to be reported violent (Eisikovits & Edleson, 1989). In this study, only 20 to 26.9 percent of the men who participated in 12-session education and combined groups were found to be violent at follow-up. Therefore, these groups achieved outcomes comparable to the best results found in prior studies that have examined violence recidivism rates.

Few studies have examined the continuing use of threats by men during follow-up. Those that have, have found a majority of men to be continuing their use of threats at follow-up (Edleson & Grusznski, 1988; Tolman, Beeman, & Mendoza, 1987). The results of this study are consistent with these previous findings. However, significant differences were found in the effects of treatment models on men's postgroup use of more severe, terroristic threats. Of the programs compared, the 12-session education groups were found to be the most effective in reducing men's postgroup use of terroristic threats.

Interpretations of the terroristic threat and threat findings will certainly vary. On the one hand, it could be argued that persistent reports of threats during follow-up mean that none of these programs was effective. On the other hand, significant differences were revealed among programs when examining terroristic threats. This finding may indicate that the educational approach is best suited to reduce severe threats but that equal attention must be directed to these less severe types of threatening behavior.

Almost 4 out of 10 of the men in this study had been ordered by the court to receive treatment. When the court variable was entered into the logit analyses for both violence and terroristic threats, it failed to prove a significant predictor. Thus, court-ordered men were as likely as any other men to respond favorably or unfavorably to the various combinations of treatment intensity and model.

Several sets of data have yet to be analyzed and will be the subject of subsequent reports. Intake-to-closing changes on several factors as well as qualitative, interview data will be examined in an effort to identify factors that differentially mediate the changes observed. Eighteen-month follow-up interviews are being completed, and these longer-term data also will be examined to determine if the changes reported in this article are maintained a year later.

Several limitations of this study deserve mention here. First, the groups studied were offered by one agency during a 12-month period, and the findings are therefore unique to this setting. Second, due to the concerns of the agency in which the study was conducted, men were assigned randomly only to treatment conditions; no-treatment control conditions were not used. It is impossible to draw conclusions from this study about the overall effects of treatment per se. Finally, as noted above, a great deal of subject attrition took place between intake and follow-up interviews. By the time the follow-up interviews were completed, the statistical power of the experiment was reduced. Only replications of this study with larger samples in a variety of agency settings can provide a strong sense of confidence that these conclusions do not reflect unique events.

These findings tend to support the greater efficiency of 12-session groups when compared with 32-session ones as offered in this study. They also tend to support the greater effectiveness of programs that are more structured and educational. Although violence and terroristic threats have been ended by most men in such programs, reports of continuing threats against women during follow-up indicate that further programmatic changes are necessary.

The results of this study should allow battered women, activists, policymakers, and practitioners to make more educated judgments concerning the efficacy of batterers' treatment programs. This study joins the ranks of a small number of other outcome studies of batterers' treatment. Many treatment programs for men who batter have

received widespread attention, received substantial funding, and worked with thousands of men. Sadly, few of these programs have made outcome data available with which battered women and the general public may judge their effectiveness. There is a clear need for greater accountability by such programs.

TABLE 1. Men Reported Violent and Nonviolent by Group Intensity of 6-month Follow-up (n = 92)[a]

Report	Number of sessions			
	12		32	
	n	%	n	%
Nonviolent	36	65.5	20	54.1
Violent	19	34.5	17	45.9

[a] Chi[sup2] = 0.77; df = 1.

TABLE 2. Men Reported Violent and Nonviolent by Treatment Model at 6-month Follow-up (n = 92)[a]

Report	Treatment Model					
	Education		Combined		Self-help	
	n	%	n	%	n	%
Nonviolent	21	67.7	23	65.7	12	46.2
Violent	10	32.3	12	34.3	14	53.8

[a] Chi[sup2] = 3.32; df = 2.

TABLE 3. Men Reported to Be Using Terroristic Threats by Group Intensity at 6-month Follow-up (n = 92)[a]

Number of sessions

Report	12		32	
	n	%	n	%
Nonthreatening	27	49.1	16	43.2
Threatening	28	50.9	21	56.8

[a] Chi[sup2] = 0.11; df = 1.

TABLE 4. Men Reported to Be Using Terroristic Threats by Treatment Model at 6-month Follow-up (n = 92)[a]

Report	Treatment Model					
	Education		Combined		Self-help	
	n	%	n	%	n	%
Nonthreatening	19	61.3	17	48.6	7	26.9
Threatening	12	38.7	18	51.4	19	73.1

[a] Chi[sup2] = 6.79; df = 2, p = .03.

TABLE 5. Logit Analysis for Continued Use of Terroristic Threats at Follow-up[a]

Term	Coefficient	Standard Error
Group models compared		
Combined and education	-.66517	.3290
Combined and self-help	.97687	.3620
Marital status at intake	.59318	.2386
Constant	.23781	.2364

Term	Approximate		
	Chi[sup2] to Remove	df	p

Group models compared			
Combined and education			
Combined and self-help	8.48	2	.0144
Marital status at intake	6.56	1	.0105
Constant	1.03	1	.3107

[a] Goodness-of-fit $\chi^2 = 0.356$, $p = .837$.

REFERENCES

- Brygger, M. P., & Edleson, J. L. (1987). The Domestic Abuse Project: A multi-systems intervention in woman battering. *Journal of Interpersonal Violence*, 3, 324-336.
- Butcher, J., & Koss, M. (1978). Research on brief and crisis-oriented therapies. In S. Garfield & A. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (2nd ea.) (pp. 725-767). New York: Wiley.
- Dobash, R. E., & Dobash, R. (1979). *Violence against wives*. New York: Free Press.
- Edleson, J. L., & Brygger, M. P. (1986). Gender differences in reporting of battering incidences. *Family Relations*, 35, 377-382.
- Edleson, J. L., & Grusznski, R. J. (1988). Treating men who batter: Four years of outcome data from the Domestic Abuse Project. *Journal of Social Service Research*, 12, 3-22.
- Eisikovits, Z. C., & Edleson, J. L. (1989). Intervening with men who batter: A critical review of the literature. *Social Service Review*, 63, 384-414.
- Ewing, W., Lindsey, M., & Pomerantz, J. (1984). *Battering: An AMEND manual for helpers*. Denver, CO: Amend.
- Feazell, C. S., Mayers, R. S., & Deschner, J. (1984). Services for men who batter: Implications for programs and policies. *Family Relations*, 33, 217-223.
- Goffman, J. M. (1980). *Batterers anonymous: Mutual support counseling for women-batterers*. Redlands, CA: Coalition for the Prevention of Abuse of Women and Children.

Gondolf, E. (1987). Evaluating programs for men who batter: Problems and prospects. *Journal of Family Violence*, 2, 95-108.

Grusznski, R. J., & Carrillo, T. P. (1988). Who completes batterer's treatment groups? An empirical investigation. *Journal of Family Violence*, 3, 141-150.

Hart, B. (1988). Safety for women: Monitoring batterers' programs. Harrisburg, PA: Pennsylvania Coalition Against Domestic Violence.

Jennings, J. L. (1987). History and issues in the treatment of battering men: A case for unstructured group therapy. *Journal of Family Violence*, 2, 193-213.

Johnson, D H., & Gelso, C. J. (1980). The effectiveness of time limits in counseling and psychotherapy: A critical review. *Counseling Psychologist*, 9, 70-83.

Jouriles, E. N., & O'Leary, K. D. (1985). Interspousal reliability of reports of marital violence. *Journal of Consulting and Clinical Psychology*, 53, 419-421

Pence, E., & Paymar, M. (1986) Power and control: Tactics of men who batter. Duluth, MN: Program Development.

Reid, W. J., & Shyne, A. (1969). Brief and extended casework. New York: Columbia University Press

Reilly, P., & Grusznski, R. J. (1984). A structured didactic model for men controlling family violence. *International Journal of Offender Therapy and Comparative Criminology*, 28, 223-235.

Roberts, A. (1982) A national survey of services for batterers. In M. Roy (Ed.), *The Abusive Partner* (pp. 230-243). New York: Van Nostrand Reinhold.

Rosenbaum, A. (1986). Group treatment for abusive men: Process and outcome. *Psychotherapy*, 23, 607-612.

Rosenbaum, A. (1988). Methodological issues in marital violence research. *Journal of Family Violence*, 3, 91-104

Saunders, D. G. (1984). Helping husbands who batter. *Social Casework*, 65, 347-353.

Saunders, D. G. (1988). Issues in conducting treatment research with men who batter. In G. T. Hotaling, D. Finkelhor, J. T. Kirkpatrick, & M. A. Straus (Eds.), *Coping with family violence* (pp. 145-156). Newbury Park, CA: Sage Publications.

Sherman, L. W., & Berk, R. A. (1984). The specific deterrent effects of arrest for domestic assault. *American Sociological Review*, 49, 261-272.

Sonkin, D. J., Martin, D., & Walker, L. E. A. (1985). *The male batterer: A treatment approach*. New York: Springer.

Straus, M. A. (1979). Measuring intrafamily conflict and violence: The Conflict Tactics Scale. *Journal of Marriage and the Family*, 41, 75-88.

Szinovacz, M. E. (1983). Using couple data as a methodological tool: The case of marital violence. *Journal of Marriage and the Family*, 45, 633-644.

Tolman, R. M., Beeman, S., & Mendoza, C. (1987, July). The effectiveness of a shelter-sponsored program for men who batter: Preliminary results Paper presented at the Third Family Violence Conference for Researchers, Durham, NH.

Tolman, R. M., & Edleson, J. L. (1989). Cognitive-behavioral intervention with men who batter. In B. A. Thyer (Ed.), *Behavioral family therapy* (pp. 169-190). Springfield, IL: Charles C. Thomas

Videka-Sherman, L. (1988). Metaanalysis of research on social work practice in mental health. *Social Work*, 33, 325-337

Walker, L. E. (1979). *The battered woman* New York: Harper & Row.

Accepted October 19, 1989

~~~~~

Jeffrey L. Edleson and Maryann Syers

Jeffrey L. Edleson, PhD, is Professor, School of Social Work, University of Minnesota, Minneapolis, and Director of Evaluation and Research, Domestic Abuse Project, 204 West Franklin Avenue, Minneapolis, MN 55404. Maryann Syers, PhD, is Instructor,



School of Social Work, University of Minnesota, and Research Associate, Domestic Abuse Project. The research summarized in this article was funded by grants from the Bush Foundation of St. Paul, MN. The authors wish to acknowledge the staff and clients of the Domestic Abuse Project, without whom this study would not have been possible.

---

Copyright of Social Work Research & Abstracts is the property of National Association of Social Workers and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

[Result List Refine Search](#) **Result 2 of 10**

## **A TEST OF THE EFFICACY OF COURT-MANDATED COUNSELING FOR DOMESTIC VIOLENCE OFFENDERS: THE BROWARD EXPERIMENT\***

**LYNETTE FEDER\*\***

**University of Memphis**

**LAURA DUGAN\*\*\***

**University of Maryland**

This study investigated the effectiveness of a popular batterers' intervention program in reducing repeated violence among men who were convicted of misdemeanor domestic violence. A classical experimental design randomly assigned all 404 male defendants in Broward County Courthouse into an experimental (one-year probation and court-mandated counseling) or control (one-year probation only) conditions. The study followed these men for 12 months, collecting information from offenders' self-reports, victims' reports, and official measures of rearrests. No significant differences were found between the experimental and control groups in their attitudes, beliefs, and behaviors regarding domestic violence; both groups were equally likely to engage in both minor and severe partner abuse. In addition, no significant differences were found between the two groups in their rates of rearrest. Further analyses indicated that stake-in-conformity variables (employment and age) predicted both attendance at treatment and reoffending.

---

\* This research was supported by a National Institute of Justice grant award (NIJ-96-WT-NX-0008). Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice. This study would not have been possible without Judges Geoffrey Cohen and Alfred Horowitz and their deep commitment to helping victims. Thanks to Drs. Robert Boruch and Lawrence Sherman for their encouragement and advice. Sincere appreciation to a great team of research assistants—Dulce Guitierrez, Cassandra Harmon, Chelsea Marder, Debora Rothmond, Linda Ryan and Laurie Benson Schultz. Finally, we are grateful to the many victims who spoke with us candidly and at length. Address all correspondence to Lynette Feder at [lynettedfeder@juno.com](mailto:lynettedfeder@juno.com).

\*\* Lynette Feder is an associate professor at the Department of Criminology and Criminal Justice, University of Memphis. She received her Ph.D. from the University of Albany, School of Criminal Justice. Her recent research interests are directed at rigorously evaluating criminal justice interventions to address both policy questions and underlying theoretical issues.

\*\*\* Laura Dugan is an assistant professor at the Department of Criminology and Criminal Justice at the University of Maryland, College Park, and is a member of the National Consortium on Violence Research. She received her Ph.D. in public policy and management (1999) her M.S. in statistics (1998), and her M.S. in public management and policy (1995) from Carnegie Mellon University. She investigates the consequences of criminal victimization and the efficacy of victimization prevention policy and practice. She also examines the effects of domestic violence laws, policies, services, and reductions in welfare benefits on the behavioral patterns of offenders, victims, and criminal justice participants.

JUSTICE QUARTERLY, Vol. 19 No. 2, June 2002

© 2002 Academy of Criminal Justice Sciences

In the past 20 years, numerous policies, procedures, and programs have been instituted, aimed at reducing or curtailing domestic violence in the United States. In the 1980s, there was a rapid growth in pro-arrest policies for domestic violence (Sherman, 1992). With increased arrest rates, pressure was placed on the courts to find a way to deal with these domestic violence offenders (Feder, 1997). The result was a rise in the use of court-mandated counseling.

Ironically, the idea of counseling men was born out of the women's shelter movement, when it was realized that many of the victims returned to their abusive mates (Adams & McCormick, 1982; Healey, Smith, & O'Sullivan, 1998; Jennings, 1987). But by far the greatest growth in these spouse abuse abatement programs (SAAPs) was brought about by the rise in pro-arrest laws in the late 1980s (Gondolf, 1997; Hotelling & Sugarman, 1986; Johnson & Kanzler, 1993). Given this population's high rates of attrition from counseling programs, court-mandated counseling was viewed as one method of ensuring greater compliance with treatment programs. But it served other functions as well. Specifically, it provided the courts with an alternative to incarceration (Dutton, 1984; Dutton & McGregor, 1991; Harrell, 1991), especially important during a time of overcrowded jails. In addition, given overloaded court dockets, mandated counseling offered the promise of shortening court proceedings (Gondolf, 1991) and simultaneously adding to the deterrent effects of arrest (Dutton, 1987). All this was done while hope was held out of changing the behavior of domestic violence offenders and, in that way, ending the cycle of violence (Farley & Magill, 1988).

Since the earliest programs for batterers grew directly out of the women's shelter movement, it is not surprising that they borrowed heavily from this feminist orientation. Typically, the various programs encouraged men to confront their sexist beliefs and accept responsibility for their past abuse and taught them alternative behaviors and reactions (e.g., anger management, assertiveness, relaxation techniques, and communication skills). The most popular of these programs is a feminist cognitive psychoeducational approach, called the Domestic Abuse Intervention Project. Referred to simply as the Duluth Model, it focuses on battering as part of a range of male behaviors used to control women. The curriculum is taught in group sessions that emphasize the modification and development of cognitive techniques that batterers can use to avoid conflict.

Soon after SAAPs began to appear, studies evaluating their efficacy began to surface. In this first wave of evaluation research,

the results indicated high rates of success in reducing the frequency and/or severity of subsequent violence among this population of offenders (Deschner & McNeil, 1986; Neidig, Friedman, & Collins, 1985; Rosenfeld, 1992). A number of researchers have since noted that the successful results instead reflected methodological shortcomings inherent in these studies, rather than the programs' actual efficacy in reducing violence (Chen, Bersani, Myers, & Denton, 1989; Ford & Regoli, 1993; Gondolf, 1987a). These findings led one prominent team of researchers to lament, "After reviewing much of the research literature, what do we 'know' about the short and long-term effects of treatment on wife assault? The answer, unfortunately, is 'not much.' . . . We cannot confidently say whether 'Treatment works.' We should be well beyond that question, asking instead, 'What treatment works best on which types of client and under what conditions?' " (Hamberger & Hastings, 1993, p. 220).

We must acknowledge the real possibility that our well-intended interventions may have unintended harmful effects (Dishion, McCord, & Poulin, 1999; Petrosino, Turpin-Petrosino, & Finckenauer, 2000). To continue mandating counseling for convicted abusers necessarily means that limited resources will be diverted to such counseling programs from alternative programs for battered women and their children (Gondolf, 1987b; Tolman & Bennett, 1990). And what is even more problematic, there is a possibility that *ineffective* treatment may be more dangerous for the victim than no treatment at all. Specifically, research has indicated that the most influential predictor of an abused spouse's return to her husband is his participation in counseling (Feazell, Mayers, & Deschner, 1984; Gondolf, 1987b). If treatment is essentially ineffective in decreasing recidivism, then continuing to mandate treatment may be inadvertently providing these victims with a false sense of security that, in the end, may lead to a higher likelihood of future injury (Hamberger & Hastings, 1993; Harrell, 1991).

This study's central question was whether courts can affect change among men who are convicted of misdemeanor domestic violence by mandating a specific counseling program. We begin with a review of the empirical research on court-mandated counseling interventions. We then report the results of our study, which used an experimental design to test the efficacy of court-mandated counseling for men who were convicted of misdemeanor domestic violence in one county in south Florida.

## REVIEW OF RESEARCH

A plethora of studies have evaluated SAAPs, but the quality of these studies has varied considerably. This variable quality

presents a problem for evaluating whether mandated counseling of domestic violence offenders is effective. The National Research Council has referred to a hierarchy of research designs when analyzing results across studies. It has evaluated family violence intervention programs by organizing these studies as (in order of increasing rigor) preexperimental, quasi-experimental, and experimental designs (Chalk & King, 1998). Similarly, Sherman et al. (1997) reviewed crime prevention studies and concluded that not all studies should be given equal weight when making policy recommendations. A basic argument is that it is important to separate results based on the strength of the research designs and to attribute greater confidence in the validity and generalizability of more rigorous studies. Hamberger and Hastings (1993) reported on inconsistencies in the literature on the effectiveness of counseling. In addition, Feder and Forde (2000) demonstrated that there was a higher likelihood of finding treatment efficacy as one used increasingly weaker research designs. These findings should not surprise evaluators.

As many have noted, the classical experimental design, with random assignment to experimental and control groups, provides the most rigorous test of an intervention (Berk, Boruch, Chambers, Rossi, & Witte, 1985; Chalk & King, 1998; Farrington, 1983). Therefore, we concentrate our review on the four studies that used an experimental design to test the effectiveness of court-mandated counseling in reducing subsequent violence among men who were convicted of domestic violence. There is no doubt that each study has limitations. However, recognizing these limitations will provide a fuller understanding of the meaning of each study's results and its applicability.

*Palmer, Brown, and Barrera*

Palmer, Brown, and Barrera (1992) conducted the first experiment, done in Canada, to test the efficacy of court-mandated counseling for abusive husbands. In their experiment, court-involved men were randomly assigned to an experimental group mandated to attend counseling ( $n = 30$ ) or a control group ( $n = 29$ ) that was not so mandated. Of the men in the experimental group, 70% completed treatment, defined as attending 7 or more of the 10 sessions. Recidivism was operationalized using information from police reports of repeated abuse. According to Palmer et al., men who were mandated into counseling were significantly less likely to recidivate than men who were not (10% versus 31%, respectively).

Palmer and her colleagues noted some of the limitations of their experiment. First, they discussed the problem of relying exclusively on official statistics in testing the efficacy of court-mandated counseling. They reported that they attempted also to collect survey information from the batterers and their victims, but the information was of limited use because of low survey response rates (55% and 22%, respectively). Palmer et al. also noted that there were strong ethical concerns about the experiment at the project site. Unfortunately, they did not elaborate on the extent and implications of these concerns.

In reading the results of this study, we suggest that there are two troubling aspects that may threaten the validity of the finding concerning the efficacy of counseling. First, the researchers did not tell how this sample of 59 men was derived, other than that the sampling took place over a 17-month period. Given that Palmer et al. conducted their research in a relatively large jurisdiction and had only a small sample ( $N = 59$ ), it is probable that not all men who were convicted of wife abuse and placed on probation in this jurisdiction were necessarily included in the experiment. Since the researchers did not report their sampling frame, it is possible that the men in the sample represent only those for whom there was a consensus among court professionals that they could be involved in the experiment. Or, it may be that the batterers themselves had to consent to the treatment program and to being part of the study. In the former case, the issue is external selection bias, whereas in the latter case, the issue is self-selection. If either case is true, both alternatives would lead to a higher likelihood of finding treatment effects, since the sample may have included the "better" or "more motivated" candidates for treatment.

Second, and most troubling, Palmer and her colleagues failed to find a relationship between attendance at counseling sessions and recidivism. If treatment, per se, were responsible for decreasing the likelihood of future violence among these men, one would expect attendance at treatment to correlate with recidivism. This result could also be related to the small sample size, which limited the power of the experiment to test for an attendance effect and ultimately for the efficacy of counseling.

### *Ford and Regoli*

Ford and Regoli (1993) tested different prosecutorial responses to domestic violence by randomly assigning subjects to one of three groups: (1) pretrial diversion to a spouse abuse abatement program, (2) conviction with a recommendation to a SAAP as a condition of probation, or a (3) recommendation for presumptive sentencing not

including counseling. This study included all men who were formally charged with a misdemeanor assault against a female conjugal partner during a 13-month period who had not previously been prosecuted for a violent act against the same victim, did not have a criminal history of felony violence, or were not viewed by the prosecutor as posing a serious threat of imminent danger to the victim ( $N = 678$ ). Failure was defined as assaulting the original victim. Information was gathered from victims' reports 6 months postadjudication (a 63% response rate). Ford and Regoli found no significant differences in the rate of reassault for any of the three groups. In addition, their failure to find differences was not dependent upon whether treatment was assigned or received.

Ford and Regoli's study has greater power than Palmer et al.'s study (1992) because of its larger sample size. Furthermore, it enabled tests to be made for the effects of attending counseling. However, it has two potential limitations that may threaten its internal validity. First, though the sample size is large, Ford and Regoli did not report whether (and how many) batterers were rejected from inclusion in the sample because the prosecutor viewed them as posing "such a serious threat of imminent danger that the prosecutor took immediate action against the suspect prior to his inclusion in the experiment" (p. 150). The language of the article makes it clear that the more violent men who were convicted of misdemeanor domestic assault in this population frame were likely to be weeded out of the sample. On the basis of statistical regression to the mean (Campbell & Stanley, 1963), the decision to remove these men may have made it more difficult to find a treatment effect, since these men's violence may have been moderated in follow-up measures. Equally problematic in interpreting the findings is that the interviews with the victims occurred six months postadjudication. Since the men may not have immediately begun their counseling upon adjudication, it is not known how much treatment they had received when the posttests were conducted. Recall that Palmer et al. (1992) provided a criterion of 7–10 counseling sessions to be considered a treatment completer. Ford and Regoli did not report such a benchmark in their study.

#### *Davis, Taylor, and Maxwell*

The Brooklyn (Kings County Criminal Court) Experiment (Davis, Taylor, & Maxwell, 2000) tested the efficacy of court-mandated counseling for men who were convicted of misdemeanor domestic violence. The 376 men were randomly assigned to a 40-hour spouse abuse abatement program that was based on the Duluth Model (offered in a 26-week format and later in an 8-week format) or to a

control condition of 40 hours of community service. To be included in the study, the judge, prosecutor, and defendant all had to agree to treatment for the defendant. Official records and surveys of batterers and victims were used to track differences in the three groups' performance. Davis et al. reported that the men in the 26-week program did significantly better in terms of new criminal justice incidents than did the men in either the short treatment program or the control group (10%, 25%, and 26%, respectively). They also suggested that the results from interviews with victims (although not statistically significant) were indicative of a similar pattern, with the victims reporting fewer new incidents for men in the long program than in the short program or community service (15%, 18%, and 22%, respectively). Davis et al. concluded, "Taken together, these studies provide a case for rejecting the null hypothesis that treatment has no effect on violent behavior toward spouses" (p. 76).

As with the other experiments, the Brooklyn Experiment has a few limitations that need to be noted. First, Davis et al.'s population frame included all men who were adjudicated for misdemeanor domestic violence in Brooklyn during a one-year period for whom the judge, prosecutor, and defendant agreed to treatment. Though more than 11,000 domestic violence cases were adjudicated in that period, there were only 376 cases in which all the parties agreed to the treatment program. Thus, there are potential threats to the internal and external validity of Davis et al.'s study because these individuals may not be representative of all batterers. By including only men who were willing to be counseled, a selection bias was introduced. The results of this selection bias may have made it more likely to find a treatment effect, since those who were not interested in treatment were not included in the experiment. This selection bias also poses a potential threat to external validity in that the results from this experiment may not apply to what occurs outside the experiment. Most jurisdictions that have adopted court-mandated counseling programs have not set such a threshold.

More problematic, though, is the researchers' assumption that 40 hours of community service is an adequate control condition for a court-mandated intervention program for batterers. That is, Davis et al. failed to provide information to demonstrate that men who were mandated into the 40-hour community service program versus those who were mandated to attend the treatment program were treated with the same degree and duration of criminal justice supervision and monitoring. Without this information, it does not



seem reasonable to assume that the groups being compared received the same things in all ways except that one group was mandated to receive counseling. This point is critical in assessing the internal validity of Davis et al.'s experiment.

In fact, the results fit an alternative interpretation that focuses not on treatment but on criminal justice monitoring. It is easy to imagine that men who were mandated into a 40-hour community service program may have completed that judicial sanction much more quickly than those who were mandated to attend a 26-week intervention program. If criminal justice oversight realistically ended at the point when the men met the requirements of their sanction, then those who were mandated into the 26-week treatment program would have received more criminal justice supervision than those who were told to attend the 8-week batterers' program who received more monitoring than those who were given a 40-hour community service sanction.

The differences that Davis et al. observed among the groups may therefore be due to differences in supervision, rather than differences in treatment. This interpretation fits the results presented by the researchers. First, there was no indication that treatment led to changes in batterers' attitudes about domestic violence, conflict resolution, or responsibility for their actions. Second, more men completed the 8-week batterers' program in comparison to the 26-week program. If treatment, not supervision, was responsible for differences among the groups on arrest outcomes, then one would expect that the men who were assigned to the shorter (more-likely-to-be-completed) program would have performed as well or better than those who were mandated into the longer program. Logic dictates that if treatment is effective, then attendance should relate to recidivism. Yet, the men in the longer counseling program performed consistently better than the two other groups. In addition, the performance of those who were ordered into the shorter counseling program was more similar to those who were mandated into the community service (control) group.

### *Dunford*

Dunford (2000) completed the largest experiment, to date, on the efficacy of court-mandated counseling. Working out of a naval base, the study randomly assigned counseling to all men ( $N = 861$ ) who had a substantiated physical assault against their wives. Batterers' and victims' reports and official measures were examined over a one-year period. Dunford reported in his analysis of victims' reports and official measures of rearrest that there were no significant differences in outcomes between the control and experimental

groups. The direct implication of his findings is that mandated counseling was ineffective for a navy population.

It is important to note that Dunford had the full cooperation of the navy, which gave him more control over the implementation of this experiment than would typically be afforded in a social science setting (for a discussion, see Berk, Smyth, & Sherman, 1988; Feder, Jolin, & Feyerharm, 2000; Petersilia, 1989). Dunford was able to ensure high treatment fidelity with a low misassignment rate, and his study had a high victim response rate. That is, fidelity may have been higher because men in the navy may be more likely to attend treatment than are men from the general population. Misassignment was low because of close oversight and cooperation of the navy with the researcher. And, the study had a higher-than-usual response rate on the victims' survey because women in the navy community were much easier to locate than were women in the general population. Although Dunford's study undoubtedly has the highest internal validity of the four experiments that have been presented, there is one important concern regarding its external validity. Specifically, is it valid to generalize from enlisted men living on a naval base to men in the general population who do not live in such a structured setting?

#### *Limitations of the Four Experiments*

These four experiments have provided some of the strongest evidence testing the efficacy of mandated counseling for reducing future episodes of domestic violence. The larger literature provides further suggestive evaluations about the efficacy of mandated counseling (Feder & Forde, 2000). A comparison of the results of these experiments, however, shows that there are inconsistencies and that questions remain about the integrity of these experiments and the generalizability of their results.

In sum, we suggest that there is a need to address three key issues when implementing an experimental design. First, a study must include men from the general population, rather than from a smaller specified subpopulation, so the results may be more generalizable to jurisdictions throughout the country. Second, in testing the efficacy of mandated counseling, it is important that the experiment not be limited to men for whom they and/or others agree to a batterers' treatment program. Not limiting the sample population in this manner will ensure that the study results will be more generalizable to other jurisdictions that implement court-mandated counseling programs. Third, it is critical to the integrity of the design that participants receive the same things in the same amount—other than the experimental stimulus—throughout the

study. Specifically, men in both experimental and control conditions would have to receive the same amount and duration of criminal justice supervision, monitoring, and nontreatment contact so that the only differences between the two groups would be that one group was court mandated into the counseling program and the other was not.

Only when these steps have been taken can researchers address the important policy question, "Can courts effect change in spousal assault by mandating men who are convicted of misdemeanor domestic violence into a spouse abuse abatement program?" This question dictates both who is to be studied and how. Specifically, a study must include all men who were mandated into this counseling program, whether or not they attend and/or complete this treatment. Since some subjects may not be mandated into counseling but may go on their own or be assigned to counseling and still not attend, the study must be able to analyze for treatment assigned as well as treatment received.

## METHODS

### *Experimental Design*

The Broward Experiment, which used a classical experimental design, took place in Broward County (an area encompassing Fort Lauderdale, Florida) in the two courts exclusively charged with handling domestic violence cases. At the time of the experiment, the police responded to domestic violence calls using a presumption-in-favor-of-arrest response. During a five-month period in 1997, all men who were convicted of misdemeanor domestic violence<sup>1</sup> in this jurisdiction (and the women who were the victims of these assaults) were placed in the study. The only exceptions were when either the defendant or victim did not speak English or Spanish; either the defendant or victim was not aged 18 or older; the defendant was severely mentally ill; or, at the time of sentencing, the judge allowed the defendant to move to another jurisdiction and serve his probation through mail contact. All other defendants ( $n = 404$ ) were included in the study.

---

<sup>1</sup> Note that the terms *convicted* and *adjudicated* have legal significance. The population we used included men who had either (1) pled guilty or no contest to charges of domestic violence battery or who were found guilty after trial and were placed on probation; (2) were placed on probation, whether adjudicated guilty or not, for the offense of domestic violence battery; or (3) were found guilty of or placed on probation for crimes of domestic violence. In fact, the vast majority of defendants (96%) pled no contest to the charges. For ease of use, throughout this report we refer to this entire group of men as those who were adjudicated or convicted of misdemeanor domestic violence charges.

Men were randomly assigned to an experimental or a control group. Men in the experimental group were sentenced to one-year probation and 26 weeks of group sessions from a local batterers' intervention program. Men in the control group were sentenced to one-year probation. Random assignment was based on the last digit of the court docket number assigned to the case—randomly generated by the court computer: the experimental (even number) or the control group (odd). This method allowed the judges to implement the random assignment process quickly. Just as critically, it allowed the researchers to identify departures from the random assignment process. At sentencing, the judges referred men from the experimental group to one of five county batterers' treatment programs. Each program used a 26-week program following the Duluth Model. All had previously been county certified as running an acceptable domestic violence intervention program, in line with the Duluth Model. As with all conditions of probation, the county's probation office was charged with monitoring each man's progress in meeting the judge's sentence.

#### *Data Collection*

In an effort to capture the true amount of change in individuals undergoing court-mandated counseling, the Broward Experiment gathered information from batterers, victims, and official records. Individuals did not have to volunteer to be part of the experimental protocol, but they did have to consent to be interviewed. The batterers were interviewed at the time of adjudication and 6 months postadjudication. The victims were interviewed at adjudication and at 6 and 12 months postadjudication. Official records were collected from the men's probation records about their demographic and criminal history information. We also used probation records and computer checks with the local police to track new arrests of the defendants for one-year postadjudication.

#### *Attendance at Domestic Violence Counseling*

The vast majority of men (97%) who were placed in the experimental group were assigned to attend 26 group-counseling sessions over a 26-week period. If a man missed a session, he was required to make it up. Almost 29% attended all the sessions without missing any of them, and approximately 95% missed five or fewer sessions. Eventually, approximately two-thirds attended all the sessions, and about 13% attended none. Of the men in the control group who were not mandated to attend counseling, 97% attended no classes while 3% ( $n = 5$ ) did. Measures associated with counseling included a count of the number of classes attended and a

dummy variable on whether counseling was mandated (1 = experimental group) or not (0 = control group).

### *Dosage of Domestic Violence Counseling*

Our analysis examined differences between the groups at the time of adjudication (Time 1), at six months postadjudication (Time 2), and changes between Time 1 and Time 2. Since offenders do not necessarily begin counseling immediately at adjudication, it was important to assess the amount of SAAP sessions they had completed at Time 2. We found that 30% ( $n = 35$ ) of the batterers had concluded their entire counseling program at the time of the second interview. More important for purposes of validity regarding the effects of counseling, these men had completed an average of 22 of the 26 mandated counseling sessions—or approximately 85% of the mandated dosage of counseling—at Time 2.

### *Batterers' Profile*

Information about the men in the sample was drawn from the offenders' survey, probation records, and police records completed at adjudication. The men's mean age was 35 years ( $SD = 10$ ), and their age range was from 19 to 71 years. The men described their living arrangements with the women at the time of the offense as married (49%), girlfriend/living together (38%), girlfriend/not living together (7%) and ex-girlfriend or ex-wife (7%). Residential stability was high, with the men averaging 160 months as county residents ( $SD = 142$ ). Several comparisons of this sample can be made to the 1990 U.S. census estimates (see Broward County Government, n.d.). Relative to the county, whites were underrepresented (57% versus 75%), blacks were overrepresented (36% versus 15%), and the Hispanic representation was nearly equivalent to the number in the south Florida population (6% versus 9%).

The batterers in our sample spanned all socioeconomic groups, but, on average, they tended to be of lower socioeconomic status than the averages present in Broward County and the United States. In addition, these men's educational levels were lower than the average in Broward County or nationally, since only 9% of the men in our sample said they had graduated college (compared to 19% in Broward County and 20% in the United States). Furthermore, 72% reported being employed at the time of adjudication, with most of them being at their place of employment for two or fewer years. The mean annual personal income was also lower for our sample (\$20,688 versus \$27,569 in Broward County), which probably reflects the large percentage of men in our sample (47%)

who reported that they worked in unskilled or semiskilled positions. Finally, home ownership was also low: 67% reported renting, and 33% reported owning their homes.

Probation reports, as well as the presentence investigation reports, provided information on the men's criminal histories and indicated that a large number of the men had criminal records prior to this instant offense. Forty percent of the men had one or more misdemeanor arrests (averaging about 0.9 misdemeanor offenses per individual), and 20% had one or more felony arrest (averaging 0.3 prior felony arrests per offender). In addition, 44% of the men had one or more jail stays, and 7% had one or more imprisonments. It is interesting that for 85% of the men, this was their first arrest for domestic violence.

With regard to the instant offense, police reports noted that approximately 28% of the incidents involved alcohol and another 3% involved drugs. They also noted victims' injuries in 74% of the cases. Of the incidents in which injuries were noted, the injuries consisted most often of black-and-blue marks (58%), although 8% were severe enough to require hospitalization of the victims. The men were taken into custody 99% of the time.

### *Victims' Profile*

The victims' profile is based on responses to the victims' survey at Time 1. The average age of the women was 34 years ( $SD = 9$ ), with a range of from 18 to 63 years. Thus, the women were two years younger than the men ( $SD = 6$ ), with differences ranging from 23 years younger to 14 years older. These women described the offenders as their husbands (53%), boyfriends/living together (37%), boyfriends/not living together (5%), and ex-boyfriends or ex-husbands (5%). They reported the average length of their relationship with the defendants as 7 years ( $SD = 7$ ).

The women also tended to be of a lower socioeconomic status than the averages in Broward County and nationally. About 10% had graduated from college, 47% were employed full time, 19% were employed part time, 11% were homemakers, and approximately 3% were unemployed and looking for a job. Of those who were working, 63% reported they were in unskilled or semiskilled positions, and almost 20% reported that they were in professional or managerial positions. This finding suggests that the sample of victims included the range of women in the sample. However, we suspect that women of higher occupational status may be over-represented in our victim sample because 90% of the women reported that their husbands or boyfriends were working, while only 72% of the men reported that they were employed full- or part time.

## EXPERIMENTAL INTEGRITY

### *Outcome of Random Assignment Process*

Before we turn to the findings, we first look at whether the experiment was successfully implemented, thereby ensuring internal validity when interpreting the results. A one-sample  $t$ -test of all the preassignments to the two groups based on the last digit of the court docket number shows that 238 men (53.4%) were assigned to the experimental group and 208 men (46.6%) were assigned to the control group. This split of cases does not differ from chance ( $t = 1.40, p > .05$ ). However, 42 of these cases were dropped from the sample because they failed to meet the criteria for inclusion listed earlier. Thus, there were 216 men (53.5%) in the experimental group and 188 men (46.5%) in the control group. This result again was not significantly different from chance ( $t = 1.42, p > .05$ ). Therefore, the court docket number was indeed randomly generated. However, in 14 cases (3.5%), the judges overrode the random assignment, placing men who were assigned to the control (no treatment) group into the treatment group. A total of 174 men were placed in the control group (43%), and 230 men were placed in the experimental condition (57%). With the misassignments added in, the likelihood of a split of this magnitude is low ( $t = 2.81, p < .01$ ). Nonetheless, the rate of misassignment (3.5%) is one of the lowest and well within the 0 to 10% range noted by Boruch (1997) as acceptable when implementing an experiment.

### *Tests for Equivalence of Experimental and Control Groups*

On the basis of information from the probation folders, we tested whether the men in the experimental and control groups were equivalent at the time of adjudication. If randomization was successful, then these groups should have been the same before the introduction of the experimental intervention. We compared information from the probation intake forms (completed when the defendants were adjudicated), as well as information that was available from the court records. As Table 1 indicates, the groups were statistically equivalent on all but one variable associated with criminal record, instant incident, stake in conformity, and offender's demographics. The one exception was offender's age. Specifically, the control group was two years younger than the experimental group. Age is a factor that has consistently been found to be negatively related to spouse abuse and recidivism in the research literature on domestic violence (Edleson, Eisikovits, & Guttman, 1985; Hamberger & Hastings, 1990; Hotaling &

**Table 1. Comparison of Offenders in Experimental and Control Groups**

| Demographics                               | Control Group<br>( <i>n</i> = 174) |           |                | Experimental Group<br>( <i>n</i> = 230) |           |                |
|--------------------------------------------|------------------------------------|-----------|----------------|-----------------------------------------|-----------|----------------|
|                                            | % or Mean                          | <i>SD</i> | Valid <i>N</i> | % or Mean                               | <i>SD</i> | Valid <i>N</i> |
| <i>Age (in years)**</i>                    | 33.9                               | 8.5       | 173            | 35.9                                    | 10.3      | 229            |
| <i>Race</i>                                |                                    |           |                |                                         |           |                |
| White                                      | 57.3                               |           | 98             | 56.0                                    |           | 126            |
| Black                                      | 33.9                               |           | 58             | 37.3                                    |           | 84             |
| Hispanic/other                             | 8.8                                |           | 15             | 6.7                                     |           | 15             |
| <i>Stake in Conformity</i>                 |                                    |           |                |                                         |           |                |
| Married (% yes)                            | 47.1                               |           | 170            | 42.8                                    |           | 222            |
| County resident (months)                   | 160                                | 150       | 139            | 149                                     | 137       | 193            |
| State resident (months)                    | 198                                | 158       | 134            | 180                                     | 152       | 179            |
| Home ownership (% yes)                     | 31.9                               |           | 138            | 34.1                                    |           | 182            |
| Education (0-24 scale)                     | 12.3                               | 2.2       | 144            | 12.0                                    | 2.3       | 190            |
| Employed (% yes)                           | 75.6                               |           | 165            | 68.9                                    |           | 219            |
| <i>Type of work</i>                        |                                    |           |                |                                         |           |                |
| Service                                    | 11.4                               |           | 13             | 12.9                                    |           | 18             |
| Laborer-unskilled                          | 14.0                               |           | 16             | 15.8                                    |           | 22             |
| Operative-semiskilled                      | 20.2                               |           | 23             | 18.7                                    |           | 26             |
| Craft worker-skilled                       | 22.8                               |           | 26             | 20.1                                    |           | 28             |
| Office and clerical                        | 10.5                               |           | 12             | 8.6                                     |           | 12             |
| Sales                                      | 7.0                                |           | 8              | 11.5                                    |           | 16             |
| Technician                                 | 1.8                                |           | 2              | 0.7                                     |           | 1              |
| Professional                               | 2.6                                |           | 3              | 5.8                                     |           | 8              |
| Official and manager                       | 9.6                                |           | 11             | 5.8                                     |           | 8              |
| Employed (in months)                       | 36.9                               | 48.9      | 19             | 91.3                                    | 192.3     | 30             |
| Income (monthly)                           | 1,670                              | 1,237     | 79             | 1,771                                   | 1,350     | 90             |
| <i>Prior Criminal Record</i>               |                                    |           |                |                                         |           |                |
| Arrest: Felony (number)                    | .30                                | .98       | 157            | .52                                     | 1.54      | 219            |
| Arrest: Misdemeanor (number)               | .94                                | 1.94      | 158            | 1.01                                    | 2.14      | 218            |
| Conviction: Felony (number)                | .22                                | .89       | 157            | .27                                     | .82       | 217            |
| Conviction: Misdemeanor (number)           | .57                                | 1.41      | 157            | .45                                     | 1.18      | 215            |
| Juvenile arrests (% yes)                   | .02                                | .18       | 154            | .01                                     | .12       | 216            |
| Prior prison sentences (number)            | .14                                | .55       | 155            | .10                                     | .45       | 216            |
| Prior jail sentences (number)              | 1.01                               | 1.88      | 155            | 1.09                                    | 1.67      | 213            |
| First arrest for domestic violence (% yes) | 84.08                              |           | 157            | 85.92                                   |           | 213            |
| <i>Incident related</i>                    |                                    |           |                |                                         |           |                |
| Alcohol noted (% yes) *                    | 22.6                               |           | 124            | 32.3                                    |           | 155            |
| Drugs noted (% yes)                        | 2.5                                |           | 120            | 4.2                                     |           | 144            |
| Victim injuries (% yes)                    | 72.8                               |           | 147            | 74.5                                    |           | 192            |
| Black-and-blue marks (% yes)               | 57.9                               |           | 107            | 48.3                                    |           | 145            |
| Cuts and bruises (stitches) (% yes)        | 4.7                                |           | 107            | 2.8                                     |           | 145            |
| Cuts and bruises (no stitches) (% yes)     | 48.6                               |           | 107            | 54.2                                    |           | 144            |
| Burns (% yes)                              | 0.9                                |           | 107            | 0.0                                     |           | 145            |
| Gunshot (% yes)                            | 0.0                                |           | 107            | 0.7                                     |           | 145            |
| Broken bones (% yes)                       | 2.8                                |           | 107            | 0.7                                     |           | 145            |
| Other injuries (% yes)                     | 6.5                                |           | 107            | 11.7                                    |           | 145            |
| Hospitalization (% yes)                    | 7.9                                |           | 107            | 6.3                                     |           | 145            |

\*  $p < .10$ , \*\*  $p < .05$ .



Sugarman, 1986). Our finding that the control group was significantly younger than the experimental group should therefore make it easier to find treatment effectiveness.

### *Integrity of Experimental and Control Conditions*

Again, because so much is beyond the researcher's control when implementing an experimental design outside a laboratory setting, it is imperative to address the separate issue of the integrity of the experiment as implemented. We have previously discussed the low misassignment rate and the finding that the two groups were equivalent on all variables save age. Both these findings speak to the random assignment being successful in ensuring that the experimental and control groups were comparable prior to the intervention.

However, a common problem in experiments is that people may change the control condition (Petersilia, 1989). Specifically, those involved in the experiment may "compensate" by providing the control group with something additional to make up for their not receiving the experimental intervention (see Babbie, 1998). A potential threat to internal validity could have come from the judges ordering additional non-SAAP programs (typically alcohol or drug programs) that would increase monitoring or supervision for men placed in the control (no treatment) condition. The probation office could also have altered the level of supervision of either group.

Thus, we compared judicial orders for men in the experimental and control groups to see if there were differences in the judges' assignments of non-SAAP programs. The results indicated that the judges equivalently assigned alcohol and drug evaluations, supervision, and treatment programs to the men in the experimental and control groups.

We also investigated whether the probation office increased its monitoring and supervision of the men in the control group to compensate for their not receiving the spouse abuse abatement program. We compared the two groups on (1) the number of months they were not reporting to the probation officers and their probation was not violated; (2) the number of probation meetings scheduled, missed, and rescheduled; (3) the number of months there were written monthly reports for each probationer; and (4) whether the men were ordered by their probation officers to have alcohol or drug testing and the number of times they were tested. The results showed that there were no significant differences between the groups. Therefore, there is no reason to conclude that probation officers treated the two groups differently.

Last, there was potential for an alternative possibility, specifically, that the probation office may not have sufficiently monitored the SAAP attendance of the men who were placed in the experimental condition. If the men were not sufficiently sanctioned for failing to comply with the court-ordered treatment, then this experiment would not offer a true test of the efficacy of court-mandated counseling. (In the language of the debate on rehabilitation, "We never tried it, and it never worked.") We tested for this possibility by looking at the men's history of attending treatment. When men failed to attend any batterers' intervention sessions, 89% ( $n = 70$ ) were violated on one or more occasions. Furthermore, of the 9 men (11%) who did not attend all 26 sessions and who were not violated, 4 missed only one session and 1 missed only two sessions. These results suggest that the probation office adequately monitored and sanctioned these men's compliance with the court-mandated SAAP.

## DEPENDENT VARIABLES

### *Offenders' and Victims' Surveys*

A battery of standardized scales was used in the offenders' and victims' interviews to assess the outcomes of the experimental intervention. A parallel interview was developed for the victims, asking them about the batterers' behavior, responsibility, and their perception of the likelihood that they would face another physical incident within the next year. The study design asked the offenders about self-reported partner abuse at the time of adjudication and six months postadjudication. Eight items from the Crowne-Marlowe Social Desirability Scale were included to assess the extent to which offenders might be less than truthful in their responses. The victims' surveys were to be administered at the time of adjudication, six months, and one year postadjudication.

If mandatory counseling was effective, we would expect differences in the attitudes of the two groups when the men in the experimental group completed their counseling. We assessed the following dependent measures for differences between the experimental and control groups and for differences between the two groups over time.

1. An abbreviated version of the Inventory of Beliefs About Wife Beating Scale was used by pulling 11 items from 30 (Saunders, Lynch, Grayson, & Linz, 1987). This scale assesses the respondent's view of the appropriateness of wife battering and the correctness of the government intervening when such incidents come to light. We analyzed the means for the scale, with possible

values ranging from wife beating being viewed as acceptable (1) to wholly unacceptable (0) behavior.

2. A shortened six-item Attitudes Toward Women scale measured the men's perceptions of the appropriate roles for women, scaled from traditional to liberal (Spence, Helmreich, & Strapp, 1973). These perceptions were coded for analysis as agree (1) and disagree (0), with agreement indicating greater support for liberal roles.

3. Criminalization of domestic violence was assessed by asking the offenders and victims whether the domestic violence incident that brought them to court should be treated as a crime. Response categories were coded as yes (1) or no (0).

4. Attitudes about the partner's responsibility for the instant offense were measured by asking the respondents to rate their partners' responsibility for the occurrence of the instant incident. Response categories were not at all responsible (1), somewhat responsible (2), equally responsible (3), and completely responsible (4). We analyzed this variable, collapsing somewhat responsible and equally responsible into a single category.

5. Self-reported likelihood to hit their partners again was operationalized by asking the men to rate the likelihood that they would hit their partners again within the next year (Harrell, 1991). The women were asked about the likelihood that their partners would hit them again. Response categories were given in increments of 10% and labeled with anchors of no chance (0%), 50/50 chance (50%), and sure to happen (100%).

6. The Conflict Tactics Scale (hereafter CTS2), as revised by Straus, Hamby, Boney-McCoy, and Sugarman (1996), was used as a measure of verbal, physical, and sexual abuse in a relationship in the offenders' and victims' surveys. We estimated average responses for negotiation, psychological, physical, sexual coercion, and sexual abuse. Response categories were coded using a normalized scale of never (0), 1 (1), 2-5 (2), and 6 or more times (3) in the past six months.

#### *Attrition Analyses of the Data*

Attrition analyses were conducted to test for differential response rates across the experimental and control groups. Response rates for the defendants were 80% ( $n = 321$ ) for the first survey and 50% ( $n = 203$ ) for the second at 6 months postadjudication. Sample attrition analysis indicated equivalent response rates for the men in the experimental and control groups. Survey completion rates for the victims were 49% ( $n = 199$ ) for the first interview and 30% ( $n$

=122) for the second at 6 months.<sup>2</sup> A high percentage of the victims' nonresponses was due to problems in tracking the victims, whereas a high percentage of the defendants' nonresponses was due to the defendants' refusal to be interviewed. Survey response rates, such as these, are typical when working with victims of domestic violence. Undoubtedly, a few researchers have been able to establish high rates of victim retention within domestic violence research. Sullivan, Rumpitz, Campbell, Eby, and Davidson (1996), using a specific and highly labor-intensive method of tracking that they developed, were able to follow 94% of the 141 victims of domestic violence in their study over 12 months. Alternately, Dunford (2000), using domestic violence victims residing on a naval base, was able to retain 78% of the women over a one-year follow-up period. However, studies indicating low retention rates when tracking difficult-to-follow subjects longitudinally are much more typical (Orwin, Sonnefeld, Garrison-Mogren, & Smith, 1994). For instance, Dobash, Dobash, Cavanagh, and Lewis (1996) reported a 41% victim response rate at the 3-month follow-up and a 22% rate at the end of 12 months, and Palmer et al. (1992) reported a 22% response rate at 16–18 months posttreatment.

Although some researchers have obtained higher victim response rates, the Broward Experiment encountered resistance to the study, which interfered with access to victims. Agency personnel who were antagonistic to the study were opposed to the experiment because they believed that randomly assigning men to the control (no treatment) group placed the victims at an increased risk of violence. While the first author saw this as an assumption in need of testing, others viewed it as so plainly self-evident that there was no need to investigate. This type of thinking seems to be common when researchers seek to test criminal justice interventions rigorously (see Petrosino et al., 2000). Within this context, some agencies decided to withhold their cooperation, thereby making it more difficult for the first author to obtain contact information on the victims. However, under no circumstances did these agencies impede the delivery or monitoring of court-mandated services to the men who had been assigned to the experimental group, since they were fighting for *more* men to be placed in the court-mandated counseling group. (For a fuller discussion of the problems encountered in implementing this experimental study, see Feder et al., 2000.)

Although all the defendants who met the criteria were included in the sample, not all defendants and their victims consented to be

---

<sup>2</sup> A third survey of the victims, 22% ( $n = 87$ ), was also conducted.

interviewed. Sample attrition analyses of the victims' surveys comparing demographic variables were completed to test for differential response rates between the offenders and victims whose partners were assigned to the experimental or control groups. The results showed that there were no significant differences between the experimental and control groups.

#### *Official Records of Rearrest*

Probation records and computer checks with the local police were used to track new arrests of the defendants for one year post-adjudication. Information on the date and type of each subsequent rearrest was recorded. We used logistic regression to examine arrest predicted by experimental condition, attendance in the SAAP program, and stake-in-conformity variables (such as marital status and employment) as used in the Minneapolis Spouse Abuse Experiment and the Spouse Assault Replication Programs (Sherman, 1992).

## RESULTS

#### *Social Desirability*

The results of the Crowne-Marlowe Social Desirability Scale fall at a neutral level, indicating that men's responses were relatively truthful. On the basis of the men's social desirability scores at the time of sentencing (Time 1), there were no significant differences between the experimental and control groups in terms of truthfulness ( $t = -1.4$ ,  $df = 288$ ,  $p > .05$ ). The experimental group averaged 2.3 ( $SD = 0.3$ ), and the control group averaged 2.2 ( $SD = 0.3$ ) on the Crowne-Marlowe Social Desirability Scale.

#### *Offenders' Attitudes*

Table 2 reports the aggregate scores of offenders' responses to the attitudinal questions for each group and over time. Coding of agree/disagree and yes/no responses as 1 and 0 allowed us to interpret the means as approximate percentages within each group. *T*-tests were used to assess whether there were differences between groups and between groups over time.

The Inventory of Beliefs About Wife Beating showed that approximately 46% of the men viewed wife beating as acceptable behavior in various situations (see Table 2). There were no significant differences between the experimental and control groups at Time 1, Time 2, or over time. On the Attitudes Toward Women scale, these men gave neutral answers on the scale and did not show differences between groups or demonstrate changed attitudes over time. Only

**Table 2. Offenders' Attitudes by Group and Over Time**

| Variable                                | Group        | Time 1 |       |     | Time 2 |       |     | Difference |       |    |
|-----------------------------------------|--------------|--------|-------|-----|--------|-------|-----|------------|-------|----|
|                                         |              | Mean   | SD    | N   | Mean   | SD    | N   | Mean       | SD    | N  |
| Inventory of beliefs about wife beating | Control      | .46    | .04   | 126 | .41    | .13   | 90  | .04        | .14   | 80 |
|                                         | Experimental | .46    | .05   | 165 | .43    | .14   | 103 | .02        | .13   | 90 |
| Attitudes toward women                  | Control      | .49    | .02   | 128 | .41    | .16   | 90  | -.10       | .15   | 81 |
|                                         | Experimental | .50    | .06   | 166 | .42    | .18   | 103 | -.08       | .18   | 91 |
| Treat domestic violence as a crime      | Control      | .26    | .44   | 133 | .34    | .54   | 90  | .00        | .47   | 82 |
|                                         | Experimental | .27    | .45   | 174 | .34    | .48   | 100 | .06        | .46   | 92 |
| Responsibility                          | Control      | 2.09   | .56   | 132 | 2.11   | .61   | 89  | .02        | .61   | 81 |
|                                         | Experimental | 2.14   | .60   | 176 | 2.03   | .62   | 103 | -.10       | .73   | 96 |
| Likelihood of repeated abuse            | Control      | 6.72   | 17.93 | 128 | 1.88   | 8.38  | 85  | -4.53      | 17.96 | 75 |
|                                         | Experimental | 4.03   | 12.57 | 177 | 3.83   | 14.42 | 99  | 0.63       | 16.51 | 96 |

26% of the men thought that domestic violence should be treated as a crime. There is also no indication of change in the men's view of how responsible their partners were for the instant offenses that led them to court. Specifically, the men in both the experimental and control groups viewed their partners as "somewhat" to "equally" responsible for the incidents. In addition, this attitude did not show change over time.

On the likelihood of repeated abuse, the results reported in Table 2 take into account situations when the offenders no longer had contact with their wives or girlfriends. *T*-tests indicated that there were no significant differences between the two groups or between these two groups over time in the men's perceptions of the likelihood of their hitting their partners again. Note that there were wide variations in the offenders' estimates of the likelihood of hitting their wives or girlfriends at both points in time. This variation is evident in the large standard deviations relative to the size of the means.

The cognitive behavioral approach used by the Duluth Model is based on the assumption that educating men about the incorrectness of the use of verbal, physical, or sexual control over their partners will lead to changes in their beliefs and ultimately lead to changes in their behaviors. The substantive result of these analyses of men's beliefs about the legitimacy of wife beating, responsibility for these incidents, and attitudes regarding the proper roles of women is that men in the court-mandated SAAP programs had not significantly changed their attitudes compared to those in the no-treatment control group.

### *Victims' Attitudes*

The victims' attitudes toward wife abuse are reported in Table 3. The victims' interviews at Time 1 (adjudication) clearly indicated

that the vast majority of women viewed wife beating in almost all contexts as inappropriate behavior. It is not surprising that this finding runs counter to what most of the men reported. There were no differences between the victims whose partners were in the experimental or control groups, and there were no differences between the groups of victims over time. The victims also reported a more liberal view of women's roles than did their partners. *T*-tests suggested that there were no differences in the women's attitudes about the appropriate roles for women between the control and experimental groups, nor did these views significantly change over time.

**Table 3. Victims' Attitudes by Group and Over Time**

| Variable                                | Group        | Time 1 |       |     | Time 2         |       |    | Difference     |       |    |
|-----------------------------------------|--------------|--------|-------|-----|----------------|-------|----|----------------|-------|----|
|                                         |              | Mean   | SD    | N   | Mean           | SD    | N  | Mean           | SD    | N  |
| Inventory of beliefs about wife beating | Control      | .10    | .09   | 85  | — <sup>a</sup> |       |    | — <sup>a</sup> |       |    |
|                                         | Experimental | .12    | .14   | 110 | — <sup>a</sup> |       |    | — <sup>a</sup> |       |    |
| Attitudes toward women                  | Control      | .87    | .15   | 85  | — <sup>a</sup> |       |    | — <sup>a</sup> |       |    |
|                                         | Experimental | .87    | .15   | 110 | — <sup>a</sup> |       |    | — <sup>a</sup> |       |    |
| Treat domestic violence as a crime      | Control      | .58    | .50   | 86  | — <sup>a</sup> |       |    | — <sup>a</sup> |       |    |
|                                         | Experimental | .55    | .50   | 103 | — <sup>a</sup> |       |    | — <sup>a</sup> |       |    |
| Responsibility                          | Control      | 1.48   | .55   | 86  | — <sup>a</sup> |       |    | — <sup>a</sup> |       |    |
|                                         | Experimental | 1.53   | .57   | 103 | — <sup>a</sup> |       |    | — <sup>a</sup> |       |    |
| Likelihood of repeated abuse            | Control      | 19.36  | 28.43 | 86  | 9.54           | 22.45 | 52 | -7.00          | 32.00 | 47 |
|                                         | Experimental | 21.44  | 28.39 | 105 | 13.98          | 24.21 | 65 | -9.07          | 27.30 | 61 |

<sup>a</sup> This question was not included on the second survey.

The survey indicated no significant differences between the groups in the victims' perceptions of whether the instant offenses that brought them to court should be viewed as criminal. About 57% of the women, compared to 26% of the men, thought that the offenses should be viewed as crimes. There were no differences between the victims' perceptions of the instant offenses between the groups or between the groups over time.

The victims also rated their level of responsibility for the instant offenses as falling between not at all responsible (1) and somewhat responsible (2), in contrast to the men's ratings. Again, there were no significant differences in the women's perceptions of responsibility between the experimental and control groups.

Finally, no significant differences accrued at Time 1, Time 2, or in the difference over time between the two groups in the victims' perceptions of the likelihood that their partners would hit them again. Their ratings of the likelihood of this event were higher than the men's (20% versus 5%, respectively), although the victims

seemed to view the likelihood of repeated physical abuse as only a possibility.

## CTS2

*Offenders' self-reports of abuse.* Table 4 shows the men's self-reported incidents on each of the five subscales of the CTS2. We found no significant differences ( $\alpha = .05$ ) on any of the CTS2 scales between the experimental and control groups when we compared information in the offenders' surveys at Time 1 or Time 2. Additional analyses of variance (ANOVAs; means not shown) of the differences of the five scales of the CTS2 over time suggested that there were no significant differences over time. These results indicate that the offenders' self-reported abusive behavior did not change over time.

**Table 4. Offenders' Responses to the CTS2 for the Control and Experimental Groups**

| Mean (N) Scale  | Time 1        |               | Time 2       |              |
|-----------------|---------------|---------------|--------------|--------------|
|                 | Control       | Experimental  | Control      | Experimental |
| Negotiation     | 2.11<br>(117) | 2.08<br>(153) | 1.58<br>(87) | 1.68<br>(90) |
| Psychological   | .88<br>(117)  | .83<br>(155)  | .41<br>(82)  | .30<br>(93)  |
| Physical        | .27<br>(115)  | .25<br>(143)  | .06<br>(82)  | .03<br>(90)  |
| Sexual coercion | .11<br>(113)  | .10<br>(154)  | .07<br>(85)  | .04<br>(88)  |
| Injury          | .16<br>(123)  | .16<br>(159)  | .06<br>(86)  | .02<br>(94)  |

Note: Response categories were 0 = never, 1 = 1, 2 = 2-5, 3 = 6+.

*Victims' reports of partners' abuse.* Table 5 shows the results of the CTS2 for the victims. There were no differences on any of the CTS2 scales between the experimental and control groups regarding information from the survey at the time of adjudication (Time 1) or six months later (Time 2). In addition, ANOVAs of the differences of the CTS2 over time suggested that there were no significant differences over time. The women's reports are another indicator that the offenders' abusive behavior did not differ between the experimental and control groups and that it did not change over time. A comparison of Tables 4 and 5 indicates that the women reported higher levels of abuse than did the men on the physical abuse and injury subscales.



**Table 5. Comparison of Victims' Responses to CTS2 for the Control and Experimental Groups**

| Mean (N) Scale  | Time 1       |               | Time 2       |              |
|-----------------|--------------|---------------|--------------|--------------|
|                 | Control      | Experimental  | Control      | Experimental |
| Negotiation     | 1.77<br>(81) | 1.75<br>(98)  | 1.99<br>(45) | 1.86<br>(54) |
| Psychological   | 1.23<br>(79) | 1.43<br>(102) | .84<br>(45)  | 1.00<br>(56) |
| Physical        | .62<br>(80)  | .65<br>(98)   | .13<br>(42)  | .11<br>(55)  |
| Sexual coercion | .21<br>(78)  | .17<br>(97)   | .06<br>(44)  | .05<br>(57)  |
| Injury          | .37<br>(84)  | .43<br>(103)  | .02<br>(45)  | .06<br>(58)  |

Note: Response categories were 0 = never, 1 = 1, 2 = 2-5, 3 = 6+.

### *Official Measures: Rearrests*

We obtained information on the number of rearrests and type of arrests for the year the men were on probation from the probation folders and from the crime records of the county's Sheriff's Office. This information indicated that 24% of the men in both the experimental and control groups were rearrested on one or more occasions during their one-year probation.<sup>3</sup> Model 1 in Table 6 shows that at a bivariate level, there are no significant differences between the experimental and control groups. However, when we measured members of the experimental group by the number of sessions attended by each, there was a significant association (Model 2).<sup>4</sup> This finding suggests that with each additional SAAP session, the likelihood that the offender would be rearrested decreased.

Original results from the Minneapolis Spouse Abuse Experiment and its subsequent replication in five sites (the Spouse Assault Replication Programs) indicated contrary findings regarding the efficacy of an arrest in deterring repeated abuse among misdemeanor domestic violence offenders. Subsequent meta-analyses of these six studies indicated that stake-in-conformity variables could be used to explain when an intervention (in this case, arrest) was effective in reducing the likelihood of subsequent violence (Berk,

<sup>3</sup> This rate of arrest for the experimental group does not include men who were arrested for violations of their probation solely for failing to attend counseling.

<sup>4</sup> An additional .5 was added to the value of this measure for men in the experimental group to distinguish those who were court mandated to attend the SAAP sessions but failed to go to any from those who were never court mandated to attend the sessions. Also, men in the control group who chose to attend the sessions were excluded from the model.

Campbell, Klap, & Western, 1992; Sherman, 1992). Specifically, when an offender had a high stake in conformity, an arrest was more likely to deter him from future abuse.

Therefore, we next investigated the impact of the experimental intervention while controlling for stake-in-conformity variables. Information was collected from the probation folders on the batterer's employment status (proportion of months employed while on probation), his residential stability (the number of times he changed residences while under probation supervision), marital status, and age in 1997 (calculated on the basis of his year of birth). Because prior criminality is also a predictor of future arrest (Farrington, 1991), we also controlled for the number of times the batterer served a jail term prior to his involvement with this study. Data were missing on marital status, prior jail terms, employment, and residential moves while on probation. To keep these data in the model, we set the values to zero (e.g., not married, not divorced, and no jail) and then constructed dummy variables to account for any bias that might be due to this adjustment.

A man could be mandated to attend the batterers' intervention program and not attend some or all the sessions or, alternately, might have been placed in the no-treatment control group and attended counseling on his own. Five men from the control group chose to attend at least one counseling session. To distinguish the control group clearly from the experimental group, we chose to eliminate those men from the sample. We examined two measures related to the treatment intervention. The first measures assignment to the experimental group without accounting for the number of court-mandated SAAP sessions attended. The second is a more dynamic measure that accounts for the number of sessions attended.

The information necessary to conduct this analysis came from the probation folders and was available for the majority of cases (97%). The independent variables for group assignment were coded as control group (0) and experimental group (1). The average number of sessions attended for the experimental group was 19 ( $SD = 10.7$ ). Dummy variables were created to assess marital status, examining married (43.3%; coded 1) versus others (coded 0) and divorced or separated (12.4%; coded 1) versus others (coded 0). Year of birth varied from 1926 to 1978, with the average age of the defendant being 35 ( $SD = 10$ ). Also included in the regression were residential stability, employment, and prior jail terms.

The results of four logistic regression models are presented in Table 6. Models 1 and 2 show the coefficients for the bivariate logistics without additional controls. As we stated earlier, assignment to

the experimental group is not significant, while the number of session is. Models 3 and 4 include the control variables measuring stake in conformity and past criminality. The results indicate that stake in conformity, as measured by age and employment, is significantly related to rearrest, while marital status and residential stability are not. Specifically, younger men were more likely to be rearrested during their one year on probation. In addition, the proportion of months employed was found to be significantly and negatively related to the likelihood of a rearrest. In fact, the logistic regression indicated that employment was the most significant factor in this model. Since assignment to the experimental group was random, additional characteristics do not affect the result of group assignment. However, the nullification of sessions attended suggests that stake in conformity and/or prior criminality may explain why some men attended more sessions than others.

To explore this issue further, we divided the experimental group into two categories: those who attended all court-mandated SAAP sessions (compliers) and those who failed to attend all the sessions (noncompliers). This division is important because batterers who failed to attend the court-mandated assigned SAAP classes were subject to sanctions. Their failure to be deterred from violating their conditions of probation (attending the court-mandated batterers' intervention program) may be a predictor of their failure to be deterred from reoffending. Table 7 reports the coefficient estimates comparing the effects of compliers and noncompliers with that for the control group. The first column, which omits control variables, shows that men in the experimental group who attended all sessions were significantly less likely to be rearrested. By taking the exponent of this estimate (0.503), we found that the odds that compliers would be rearrested were about half that of the control group. In contrast, the odds of rearrest for men who attended fewer sessions than assigned were 2.53 times higher than those of the control group ( $\exp(0.930)$ ).<sup>5</sup>

This finding strongly suggests that men who care little about the consequences of missing their court-mandated SAAP sessions are also less concerned about the consequences of reoffending. But what distinguishes the compliers from the noncompliers? Are the compliers more motivated men who tend to follow orders? Are they men with more to lose who, therefore, have higher stakes in conformity? After controlling for stake in conformity and prior criminality, we found that the differences among the compliers,

---

<sup>5</sup> Originally, we divided the noncompliers into two groups, those who missed all the sessions and those who attended some but not all. Tests of the coefficient estimates showed no statistical difference between the effects for these two groups.

**Table 6: Logistic Regression Models Predicting Rearrest**

| Variable                                     | Coefficient Estimate SE |          |          |          |
|----------------------------------------------|-------------------------|----------|----------|----------|
|                                              | Model 1                 | Model 2  | Model 3  | Model 4  |
| <i>Batterers' Counseling</i>                 |                         |          |          |          |
| Group assigned                               | 0.056                   |          | 0.051    |          |
|                                              | 0.240                   |          | 0.272    |          |
| Sessions attended <sup>a</sup>               |                         | -0.033** |          | -0.007   |
|                                              |                         | 0.010    |          | 0.012    |
| <i>Stake in Conformity</i>                   |                         |          |          |          |
| Age                                          |                         |          | -0.038*  | -0.037*  |
|                                              |                         |          | 0.015    | 0.016    |
| Married                                      |                         |          | 0.094    | 0.130    |
|                                              |                         |          | 0.312    | 0.316    |
| Divorced or separated                        |                         |          | 0.188    | 0.182    |
|                                              |                         |          | 0.439    | 0.440    |
| Number of moves                              |                         |          | 0.148    | 0.149    |
|                                              |                         |          | 0.110    | 0.111    |
| % months employed                            |                         |          | -2.230** | -2.181** |
|                                              |                         |          | 0.423    | 0.434    |
| <i>Past Criminality</i>                      |                         |          |          |          |
| Prior jail terms                             |                         |          | 0.220**  | 0.237**  |
|                                              |                         |          | 0.071    | 0.073    |
| <i>Controls for Missing Data<sup>b</sup></i> |                         |          |          |          |
| Marital status                               |                         |          | 0.850    | 0.149    |
|                                              |                         |          | 0.756    | 0.111    |
| Probation folder                             |                         |          | 0.142    | 0.086    |
|                                              |                         |          | 0.479    | 0.508    |
| Prior jail terms                             |                         |          | 0.492    | 0.434    |
|                                              |                         |          | 0.494    | 0.514    |
| Pseudo R <sup>2</sup>                        | 0.0001                  | 0.026    | 0.162    | 0.168    |

\* =  $p \leq .05$ , \*\* =  $p \leq .01$ ; all tests are two tailed.

<sup>a</sup> A value of .5 was added to this measure for men in the experimental group. See note 5.

<sup>b</sup> Missing values were set at zero, and the control variables in this group are dummy variables for the missing values.

noncompliers, and control group disappeared (see Table 7, column 2). This finding strongly suggests that those men who attended all their SAAP sessions would have avoided rearrest even if they had not been mandated into the SAAP program. Similarly, the men who violated the judges' mandate to attend batterers' treatment are also the type of men who are more likely to recidivate. In essence, determinants that drove the men to miss their court-mandated SAAP sessions are likely the same determinants that drove them to reoffend.

To test this likelihood directly, we ran a third logistic regression using only men from the experimental group ( $n = 229$ ). Instead of predicting the estimates on rearrest, we now estimated the effects of stake in conformity and prior criminality on noncompliance. The results are listed in the third column of Table 7. As expected, the same characteristics that predict rearrest also predict missing at least one court-mandated session. In addition, variables related to stake in conformity most strongly predict noncompliance.

**Table 7: Logistic Models Predicting Rearrest and Compliance**

| Variable                                     | Coefficient Estimate <i>SE</i> |                               |                                    |
|----------------------------------------------|--------------------------------|-------------------------------|------------------------------------|
|                                              | Rearrest<br>( <i>n</i> = 395)  | Rearrest<br>( <i>n</i> = 393) | Noncompliance<br>( <i>n</i> = 229) |
| <i>Batterers' Counseling</i>                 |                                |                               |                                    |
| Compliers                                    | -0.688*                        | -0.217                        |                                    |
|                                              | 0.307                          | 0.338                         |                                    |
| Noncompliers                                 | 0.930**                        | 0.318                         |                                    |
|                                              | 0.288                          | 0.331                         |                                    |
| <i>Stake in Conformity</i>                   |                                |                               |                                    |
| Age                                          |                                | -0.035*                       | -0.052**                           |
|                                              |                                | 0.016                         | 0.019                              |
| Married                                      |                                | 0.106                         | -0.149                             |
|                                              |                                | 0.313                         | 0.392                              |
| Divorced or separated                        |                                | 0.215                         | -0.390                             |
|                                              |                                | 0.441                         | 0.607                              |
| Number of moves                              |                                | 0.139                         | 0.164                              |
|                                              |                                | 0.111                         | 0.148                              |
| % months employed                            |                                | -2.030**                      | -3.238**                           |
|                                              |                                | 0.446                         | 0.549                              |
| <i>Past Criminality</i>                      |                                |                               |                                    |
| Prior jail terms                             |                                | 0.212**                       | 0.194                              |
|                                              |                                | 0.071                         | 0.107                              |
| <i>Controls for Missing Data<sup>a</sup></i> |                                |                               |                                    |
| Marital status                               |                                | 0.805                         | 1.264                              |
|                                              |                                | 0.757                         | 1.201                              |
| Probation folder                             |                                | 0.092                         | 0.044                              |
|                                              |                                | 0.480                         | 0.635                              |
| Prior jail terms                             |                                | 0.460                         | 0.688                              |
|                                              |                                | 0.495                         | 0.794                              |
| Pseudo <i>R</i> <sup>2</sup>                 | 0.059                          | 0.167                         | 0.2774                             |

\* =  $p \leq .05$ , \*\* =  $p \leq .01$ ; all tests are two tailed.

<sup>a</sup> Missing values were set at zero, and the control variables in this group are dummy variables for the missing values.

Younger men and those who worked less were less likely to attend all batterers' sessions. Finally, there is weak evidence that men who previously served jail time were also less likely to attend all sessions ( $t = 1.82$ ).

These comparisons indicate two primary findings from our study. First, when the court mandates attendance in a SAAP program, men who do not comply and attend all their sessions are also the same men who are likely to be rearrested on a new offense; 30% of the noncompliers were rearrested compared to 13% of the compliers. However, our findings show that failure to attend all sessions of the batterer intervention program does not have a harmful effect in and of itself. It simply seems to identify the men who are more inclined to reoffend. Our second primary finding indicates the primacy of employment and youth (both viewed as stake-in-conformity variables), not attendance, in predicting rearrest among our sample of men who were convicted of misdemeanor domestic violence.

While some may assert that these findings demonstrate the need to mandate SAAP programs and then to ensure that convicted domestic violence offenders attend them, it must be remembered that, in this jurisdiction, these offenders' attendance was monitored, and the offenders were sanctioned when they failed to comply. Nevertheless, the probation office was still unable to get all the men to complete all their sessions.

## DISCUSSION AND CONCLUSIONS

The primary purpose of this study was to test the efficacy of a court-mandated standard spouse abuse abatement program for domestic violence offenders that is widely used throughout the United States. It seems clear that Broward County provided a "good test" of this court-mandated batterers' intervention program that was based on the Duluth Model. Specifically, the judges kept the misassignment rate low, referred men only to programs that were certified by the county as credible, and did not order control subjects into additional programs to increase their monitoring or supervision. Likewise, the probation office monitored the men's compliance with the mandate, and when the men failed to comply with the treatment, the office began revocation procedures.

The results show that, in this county, there were no clear and demonstrable positive effects of this court-mandated SAAP program on the offenders' attitudes, beliefs, and behaviors. An analysis of the offenders' self-reported and the victims' reports of psychological and physical abuse, using the CTS2, suggested that the offenders' behavior did not change over time. Of note, there was still evidence of physical abuse 6–12 months postsentencing. Analyses failed to uncover differences between the control and experimental subjects in their likelihood of reoffending and being rearrested during the follow-up period.

The charge to throw the full force of the law at the man who does not attend all his treatment sessions seems to beg the question. In this jurisdiction, unlike those observed by Harrell (1991) and Palmer et al. (1992), men were monitored and sanctioned. While approximately one-third of the men failed to attend the batterers' program, 100% of these men were violated for one or more conditions of probation. (And 71% of the men were specifically violated for failing to attend counseling.) In all, it seems that the probation office monitored and revoked when the men did not complete the batterers' program. Despite the monitoring and sanctions, the men still self-selected into compliers (treatment completers) and noncompliers (treatment dropouts). Additional analyses indicated that stake-in-conformity variables predicted both whether a man

would comply with his terms of probation (specifically the court mandate to attend his SAAP program) and whether he would reoffend during the follow-up period. This finding, which demonstrates the importance of stake-in-conformity variables in predicting whether a criminal justice intervention (in this case, a court-mandated batterers' intervention program) is successful in reducing recidivism in an offender population, is in line with the meta-analyses of the Minneapolis and the Spouse Assault Replication Programs.

There is no doubt that the controversy surrounding the Broward Experiment, with its random assignment of men to probation only (versus probation and counseling) groups, was an impediment to the research. It led to low response rates for victims, a high turnover of the research staff, delays, and other problems. The low response rate of victims is especially problematic, since research has consistently indicated that victims provide the best information on continuing abuse (Browning & Dutton, 1986; Dobash, Dobash, Cavanagh, & Lewis, 1998; Edleson & Bryggers, 1995). To the degree that this study collected information from multiple sources (men's self-reports, victims' reports, and official measures) and all indicated similar conclusions, we have greater confidence in the results of each separate measure. We believe that this experiment provided a valid and rigorous test of the efficacy of court-mandated counseling as presently conducted in Broward County.

In addition, we wish our study had included a larger number of men who voluntarily attended counseling, since a larger number would have allowed us to conduct additional analyses on the benefits of nonmandated counseling. As nice as that would have been, the study was probably accurate in showing that only five of the men in our sample voluntarily chose to attend these counseling sessions.

In many respects, the Broward Experiment was an exploratory study of the efficacy of court-mandated counseling. The budget was relatively small, which led to a design with strong power to assess offenders but limited resources for tracking victims. Nonetheless, the results, through triangulation, suggest that an unquestioning acceptance of domestic violence batterers' intervention needs to be challenged. There is evidence from multiple sources that it did not work in Broward County.

In conclusion, we hope that others decide to replicate this study and put this important policy question to the test in their jurisdictions so that generalizability beyond Broward County, Brooklyn, and San Diego may be assessed. We have been candid in our disclosure of the problems in this experiment in the hope that others will

learn from our mistakes in working toward strong tests of policies that can ultimately help to reduce domestic violence.

## REFERENCES

- Adams, D., & McCormick, A. (1982). Men unlearning violence: A group approach based on the collective model. In M. Roy (Ed.), *The abusive partner: An analysis of domestic battering* (pp. 170-197). New York: Van Nostrand Reinhold.
- Babbie, E. (1998). *The practice of social research*. Belmont, CA: Wadsworth.
- Berk, R., Boruch, R., Chambers, D., Rossi, P., & Witte, A. (1985). Social policy experimentation: A position paper. *Evaluation Review*, 9, 387-429.
- Berk, R., Campbell, A., Klap, R., & Western, B. (1992). The deterrent effect of arrest in incidents of domestic violence: A Bayesian analysis of four field experiments. *American Sociological Review*, 57, 698-708.
- Berk, R., Smyth, G., & Sherman, L. (1988). When random assignment fails: Some lessons learned from the Minneapolis Spouse Abuse Experiment. *Journal of Quantitative Criminology*, 4, 209-223.
- Boruch, R. (1997). *Randomized experiments for planning and evaluation*. Thousand Oaks, CA: Sage.
- Broward County Government. (n.d.). Broward County economic profile [On-line]. Available: [www.co.broward.fl.us/cni02002.htm](http://www.co.broward.fl.us/cni02002.htm)
- Browning, J., & Dutton, D. (1986). Conflict Tactics Scale: Using couple data to quantify the differential reporting effect. *Journal of Marriage and the Family*, 48, 375-379.
- Campbell, D., & Stanley, J. (1963). *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin.
- Chalk, R., & King, P. (1998). *Violence in families: Assessing prevention and treatment programs*. Washington, DC: National Academy Press.
- Chen, H., Bersani, C., Myers, S., & Denton, R. (1989). Evaluating the effectiveness of a court sponsored treatment program. *Journal of Family Violence*, 4, 309-322.
- Davis, R., Taylor, B., & Maxwell, C. (2000). *Does batterer treatment reduce violence? A randomized experiment in Brooklyn*. Washington, DC: National Institute of Justice.
- Deschner, J., & McNeil, J. (1986). Results of anger control training for battering couples. *Journal of Family Violence*, 1, 111-120.
- Dishion, T., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist*, 54, 755-764.
- Dobash, R., Dobash, R., Cavanagh, K., & Lewis, R. (1996). Re-education programmes for violent men—An evaluation. *Research Findings*, 46, 1-4.
- Dobash, R., Dobash, R., Cavanagh, K., & Lewis, R. (1998). Separate and intersecting realities: A comparison of men's and women's accounts of violence against women. *Violence Against Women*, 4, 382-414.
- Dunford, F. (2000). The San Diego Navy Experiment: An assessment of interventions for men who assault their wives. *Journal of Consulting and Clinical Psychology*, 68, 468-476.
- Dutton, D. (1984). Interventions into the problem of wife assault: Therapeutic, policy and research implications. *Canadian Journal of Behavioral Science*, 16, 281-297.
- Dutton, D. (1987). The criminal justice response to wife assault. *Law and Human Behavior*, 11, 189-206.
- Dutton, D., & McGregor, B. (1991). The symbiosis of arrest and treatment for wife assault: The case for combined intervention. In M. Steinman (Ed.), *Woman battering: Policy responses* (pp. 131-154). Cincinnati, OH: Anderson.
- Edleson, J., Eisikovits, Z., & Guttman, E. (1985). Men who batter women: A critical review of the evidence. *Journal of Family Issues*, 6, 229-247.
- Edleson, J., & Bryggers, M. (1995). Gender differences in reporting of battering incidences. In S. Stith & M. Straus (Eds.), *Understanding partner violence: Prevalence, causes, consequences and solutions* (pp. 45-50). Minneapolis, MN: National Council of Family Relations.



- Farley, D., & Magill, J. (1988). An evaluation of a group program for men who batter. *Social Work with Groups*, 11, 53-65.
- Farrington, D. (1983). Randomized experiments on crime and justice. In M. Tonry & N. Morris (Eds.), *Crime and justice: An annual review of the research* (Vol. 4, pp. 257-308). Chicago: University of Chicago Press.
- Farrington, D. (1991). Childhood aggression and adult violence: Early precursors and later-life outcomes. In D. J. Pepler & K. H. Rubin (Eds.), *The development and treatment of childhood aggression*. Hillsdale, NJ: Lawrence Erlbaum.
- Feazell, C., Mayers, R., & Deschner, J. (1984). Services for men who batter: Implications for programs and policies. *Family Relations*, 33, 217-223.
- Feder, L. (1997). Domestic violence and police response in a pro-arrest jurisdiction. *Women and Criminal Justice*, 8, 79-98.
- Feder, L., & Forde, D. R. (2000). *A Test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward Experiment* (Final report, NIJ-96-WT-NX-0008). Washington, DC: National Institute of Justice.
- Feder, L., Jolin, A., & Feyerharm, W. (2000). Lessons from two randomized experiments in criminal justice settings. *Crime and Delinquency*, 46, 380-400.
- Ford, D., & Regoli, M. J. (1993). The criminal prosecution of wife assaulters. In Z. Hilton (Ed.), *Legal responses to wife assault: Current trends and evaluation* (pp. 127-164). Newbury Park, CA: Sage.
- Gondolf, E. (1987a). Evaluating programs for men who batter: Problems and prospects. *Journal of Family Violence*, 2, 95-108.
- Gondolf, E. (1987b). Seeing through smoke and mirrors: A guide to batterer program evaluations. *Response*, 10, 16-19.
- Gondolf, E. (1991). A victim-based assessment of court-mandated counseling for batterers. *Criminal Justice Review*, 16, 214-226.
- Gondolf, E. (1997). Batterer programs: What we know and need to know. *Journal of Interpersonal Violence*, 12, 83-98.
- Hamberger, K., & Hastings, J. (1990). Recidivism following spouse abuse abatement counseling: Treatment program implications. *Violence and Victims*, 5, 157-170.
- Hamberger, K., & Hastings, J. (1993). Court-mandated treatment of men who assault their partner. In Z. Hilton (Ed.), *Legal responses to wife assault: Current trends and evaluation* (pp. 188-229). Newbury Park, CA: Sage.
- Harrell, A. (1991). *Evaluation of court-ordered treatment for domestic violence offenders: Final report*. Washington, DC: Institute for Social Analysis.
- Healey, K., Smith, C., & O'Sullivan, C. (1998). *Batterer intervention: Program approaches and criminal justice strategies*. Washington, DC: U.S. Department of Justice.
- Hotaling, G., & Sugarman, D. (1986). An analysis of risk markers in husband to wife violence: The current state of knowledge. *Violence and Victims*, 1, 101-124.
- Jennings, J. (1987). History and issues in the treatment of battering men: A case for unstructured group therapy. *Journal of Family Violence*, 2, 193-213.
- Johnson, J., & Kanzler, D. (1993). Treating domestic violence: Evaluating the effectiveness of a domestic violence diversion program. *Studies in Symbolic Interaction*, 15, 271-289.
- Neidig, P., Friedman, D., & Collins, B. (1985). Domestic conflict containment: A spouse abuse treatment program. *Social Casework*, 66, 195-204.
- Orwin, R., Sonnefeld, L. J., Garrison-Mogren, R., & Smith, N. G. (1994). Pitfalls in evaluating the effectiveness of case management programs for homeless persons: Lessons from the NIAAA Community Demonstration Program. *Evaluation Research*, 18, 153-207.
- Palmer, S., Brown, R., & Barrera, M. (1992). Group treatment program for abusive husbands: Long-term evaluation. *American Journal of Orthopsychiatry*, 62, 276-283.
- Petersilia, J. (1989). Implementing randomized experiments: Lessons from BJA's Intensive Supervision Project. *Evaluation Review*, 13, 435-458.
- Petrosino, A., Turpin-Petrosino, C., & Finckenauer, J. (2000). Well-meaning programs can have harmful effects! Lessons from experiments of programs such as Scared Straight. *Crime and Delinquency*, 46, 354-379.
- Rosenfeld, B. (1992). Court-ordered treatment of spouse abuse. *Clinical Psychology Review*, 12, 205-226.
- Saunders, D., Lynch, A., Grayson, M., & Linz, D. (1987). The Inventory of Beliefs About Wife Beating: The construction and initial validation of a measure of beliefs and attitudes. *Violence and Victims*, 2, 39-55.

- Sherman, L. (1992). *Policing domestic violence*. New York: Free Press.
- Sherman, L., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P., & Bushway, S. (1997). *Preventing crime: What works, what doesn't, what's promising?* (Report No. NCJ 165366). Washington, DC: U.S. Department of Justice.
- Spence, J., Helmreich, R., & Strapp, J. (1973). A short version of the Attitudes Toward Women Scale (AWS). *Bulletin of the Psychonomic Society*, 2, 219-220.
- Straus, M., Hamby, S., Boney-McCoy, S., & Sugarman, D. (1996). The revised Conflicts Tactics Scale (CTS2): Development and preliminary psychometric data. *Journal of Family Issues*, 17, 283-316.
- Sullivan, C., Rumpitz, M., Campbell, R., Eby, K., & Davidson, W. (1996). Retaining participants in longitudinal community research: A comprehensive protocol. *Journal of Applied Behavioral Science*, 32, 262-276.
- Tolman, R., & Bennett, L. (1990). A review of quantitative research on men who batter. *Journal of Interpersonal Violence*, 5, 87-118.

## **A meta-analytic review of court-mandated batterer intervention programs: Can courts affect abusers' behavior?**

LYNETTE FEDER

*Division of Administration of Justice, Portland State University, P.O. Box 751, Portland,  
OR 97207-4015, USA  
E-mail: lfeder@pdx.edu*

DAVID B. WILSON

*George Mason University, 10900 University Boulevard, MS 4F4, Manassas, VA 20110-2203, USA  
E-mail: dwilsonb@gmu.edu*

**Abstract.** Court-mandated batterer intervention programs are being implemented throughout the United States to address the problem of domestic violence. Prior reviews of research on the effectiveness of these programs have arrived at conflicting conclusions. This study is a systematic review of the extant research on this topic. Experimental and quasi-experimental studies that used matching or statistical controls were included. The results were mixed. The mean effect for official reports of domestic violence from experimental studies showed modest benefit, whereas the mean effect for victim reported outcomes was zero. Quasi-experimental studies using a no-treatment comparison had inconsistent findings indicating an overall small harmful effect. In contrast, quasi-experimental studies using a treatment dropout design showed a large, positive mean effect on domestic violence outcomes. We discuss the weakness of the latter design and raise concerns regarding official reports. The findings, we believe, raise doubts about the effectiveness of court-mandated batterer intervention programs.

**Key words:** batterer intervention, domestic violence, intimate partner violence, meta-analysis, recidivism

Domestic violence is defined as assaultive behavior involving adults who are married, cohabitating, or who have an ongoing or prior intimate relationship (Goolkasian 1986). Research indicates just how pervasive this problem is today. Based upon crimes reported to the police in 1998, intimate partner homicides accounted for about 11% of all murders nationwide (Rennison and Welchans 2000). The National Crime Victimization Survey (NCVS) indicated that there were about 1 million violent crimes committed against persons by their current or former spouses or significant others in 1998, with the vast majority (85%) being against female victims (Rennison and Welchans 2000).

These numbers demonstrate the extent of the problem in terms of both the amount and severity of violence that some women face. Additionally, research indicates that women who have been victims of domestic violence are at greater risk of future violence (Hilberman 1980; Hirschel and Hutchinson 1992; Langan and Innes 1986). The cost to society is enormous. Domestic crime accounts for almost 15% of the total crime costs – approximately \$67 billion per year

(Miller et al. 1996). This figure does not include the impact that domestic violence has on the children who live in these homes. Past research has established that violent homes are a risk factor for producing violent adults, thereby continuing the 'cycle of violence' (Brisson 1981; Dutton 1988; Widom 1992).

The above figures speak to the importance of finding programs that can successfully intervene with domestic violence offenders. However, individual studies evaluating court-mandated batterer intervention programs have provided very mixed findings on their effectiveness. This systematic review uses meta-analytic procedures to synthesize the extant empirical evidence on the effects that court-mandated batterer intervention programs (including pre-trial diversion programs) have, over and above the effect of routine legal interventions, on rates of recidivism.

#### *Background to court-mandated batterer intervention programs*

Decades of overlooking domestic violence as a social problem has recently been followed by an intense amount of public, private, and professional interest in this subject. One of the earliest responses to family violence was the development and growth of shelters for battered women and their children (Johnson and Kanzler 1993). Soon after their establishment, shelter staff noticed that a large percentage of abused women returned to their abusive partners (Hamberger and Hastings 1993; Jennings 1987; Snyder and Scheer 1981). Even where victims successfully separated, these men typically continued their abusive patterns with a different partner (Farley and Magill 1988; Gondolf 1987). These workers came to believe that the best way to stop domestic violence was to change the behavior of the abuser (Feazell et al. 1984).

The original focus of these programs was a direct reflection of their emergence out of the women's shelter movement. Early programs were unstructured groups working with abusive men through a combination of consciousness-raising and peer self-help provided within a context of feminist theory that spoke of men's need to control women (Adams and McCormick 1982; Johnson and Kanzler 1993). Over the next few years, batterer programs developed independently at various sites across the country. As their numbers grew, the earlier unstructured consciousness-raising groups were replaced by more structured groups using psychoeducational and/or cognitive behavioral techniques (Pirog-Good and Stets-Kealey 1985). Still, this was typically done within a feminist context (Healey et al. 1998; Jennings 1987). Most of the programs encouraged men to confront their sexist beliefs and accept responsibility for their past abuse while teaching them alternative behavioral responses like anger management, assertiveness training, relaxation techniques, and communication skills (Davis and Taylor 1999; Healey and Smith 1998).

The Domestic Abuse Intervention Project, out of Duluth, Minnesota (usually just called the Duluth Model), has emerged as one of the more prevalent and widely cited programs for treating battering men. It uses a feminist psychoeducational approach whereby men are taught that battering is part of a range of male behaviors

used to control women. To stop the battering, men are given alternative methods like time-outs, empathizing, problem-solving, and tension-reducing exercises (Pence 1983). The structured curriculum is usually offered in groups that are from 6 to 32 weeks in duration (Tolman and Edleson 1995). The Duluth Model is the model of choice for many communities with some states mandating that batterer intervention programs adhere to this model (Babcock and Taillade 2000).

Cognitive behavioral approaches have also been widely used on this offender population. Typically, batterers (usually in a group format) are offered specific tools to help them see that their acts of violence are not uncontrollable outbursts but rather predictable behavioral patterns that they can learn to stop (Healey et al. 1998). The focus in these groups is on modifying how batterers think and act by working with them on skills training and anger management techniques (Healey and Smith 1998; Tolman and Edleson 1995). Today, most treatment programs blend together aspects of psychoeducational and cognitive behavioral approaches within a feminist context in an attempt to reach a broader range of clientele (Babcock and Taillade 1999; Tolman and Edleson 1995).

Court-mandated interventions using the couple as the unit of treatment are much less widely utilized and have even been expressly prohibited in 20 states (Healey and Smith 1998). Couple counseling sees the couple as the reason for the problem and, as such, works with both members to improve communication and conflict resolution skills. This method has been criticized as blaming victims as well as potentially placing them at greater risk should they honestly express their complaints to the batterer (Babcock and Taillade 1999; Healey et al. 1998).

In 1980, California became the first state to mandate treatment for men convicted of domestic violence (Johnson and Kanzler 1993; Sonkin 1988). The greatest growth in these batterer intervention programs occurred in the late 1980s due to the rise in pro-arrest laws occurring throughout the nation (Hotaling and Sugarman 1986; Johnson and Kanzler 1993). With increasing numbers of jurisdictions presuming or mandating arrest for misdemeanor domestic violence (Dutton and McGregor 1991; Feder 1997), pressure was placed on the courts to deal with these offenders (Ford and Regoli 1993; Pence 1983). At the same time, this population was proving difficult to work with as evidenced by high rates of attrition from these treatment programs (Pirog-Good and Stets-Kealey 1985; Roberts 1982). Having the court mandate treatment was therefore viewed as one method to ensure greater compliance when treating this population (Dutton 1984; Hamberger and Hastings, 1989). Judges also saw this intervention as providing an alternative to prison (important during this period of extensive overcrowding) while simultaneously holding out the hope of breaking the cycle of violence and, in that way, truly helping victims of domestic violence.

#### *Evaluations of court-mandated domestic violence programs*

Soon after court-mandated programs began appearing, studies evaluating their effectiveness appeared. In this first wave of evaluation research, the results indicated high rates of success in reducing the frequency and/or severity of

subsequent violence amongst this offender population. However, a number of researchers noted that these findings probably reflected the methodological shortcomings of the research rather than the programs' actual effectiveness in reducing violence (Ford and Regoli 1993; Gondolf 1987). These deficiencies included small sample sizes, failure to study the total population to be evaluated (as opposed to only those who completed the program), lack of appropriate comparison groups, inadequate or variable specification of the primary outcome measures and use of unreliable measures or questionable sources of data to measure treatment outcome (Hamberger and Hastings 1993; Palmer et al. 1992; Tolman and Bennett 1990).

Since then, more rigorous research has been conducted. Unlike the earlier studies, these studies produced mixed results regarding the effectiveness of mandated batterer intervention programs in reducing violence. Over the last 15 years there have been several published reviews of the growing body of research on domestic violence interventions (Babcock et al. 2004; Cromwell and Burgess 1996; Dutton 1988; Eisikovits and Edleson 1989; Hamberger and Hastings 1993; Rosenfeld 1992; Saunders 1996; Tolman and Edleson 1995). As with the individual studies, these reviews offer mixed conclusions regarding the effectiveness of court-mandated batterer intervention programs.

Hamberger and Hastings (1993), after reviewing 28 separate studies, concluded that little is known about the short and long-term effects of these programs. Similarly, Rosenfeld concluded that "the incremental benefit of court-ordered treatment over the deterrent effects of traditional criminal justice system remedies is unclear" (Rosenfeld 1992: 205). More recently, Davis and Taylor (1999) came to a very different conclusion. Computing an average effect size across five experimental or quasi-experimental studies, they concluded that "there is fairly consistent evidence that treatment works and that the effect of treatment is substantial" (Davis and Taylor 1999: 69). Finally, Babcock et al. (2004) conducted a meta-analysis of batterer intervention programs and concluded that, "the effect size due to group battering intervention on recidivism of domestic violence is in the 'small' range" (p. 1043). "To a clinician, this means that a woman is 5% less likely to be re-assaulted by a man who was arrested, sanctioned, and went to a batterers' program than by a man who was simply arrested and sanctioned" (Babcock et al. 2004: 1004).

In 1984, the Attorney General's Task Force on Family Violence recommended court-mandated treatment as an addition to legal alternatives (U.S. Attorney General's Task Force on Family Violence 1984). Yet 20 years later, the field remains uncertain about whether these programs are more effective in reducing future violence than legal interventions alone. The National Academy of Sciences has noted that "the urgency and magnitude of the problem of family violence have caused policy makers, service providers, and advocates to take action in the absence of scientific knowledge that could inform policy and practice" (Chalk and King 1998: 2).

This study attempts to answer this call by conducting a meta-analysis using the most rigorous research on court-mandated batterer intervention programs. Recent research indicates an inverse relationship between design rigor and likelihood of

finding program effectiveness (Feder and Forde 2000; Weisburd et al. 2001). Therefore, we sought to include only the most rigorous research. Like Babcock and her associates, this included experimental designs. Unlike Babcock, we did not include all quasi-experimental studies but instead limited inclusion to those which established pre-treatment equivalence between groups, either via a matched groups design or statistical controls. We also excluded studies that compared one treatment type to another, unless it also included a no-treatment control group. Additionally, in order to be included, the study had to follow offenders for six months post-intervention and use one or more objective measures of repeat violence (i.e., official or victim reports of his continued abuse). The meta-analysis seeks to examine the effect that these court-mandated batterer interventions have on this population's recidivism rate above and beyond what would have been expected through routine legal interventions.

## Method

### *Criteria for inclusion of studies in the review*

We sought to assess the effects of post-arrest mandated interventions (including pre-trial diversion programs) in reducing domestic violence offenders' future likelihood of re-assaulting through a synthesis of the extant empirical literature. To be included in this synthesis, a study had to meet the following criteria. First, the study used an experimental or rigorous quasi-experimental design. Experimental designs were defined as those using random assignment to treatment and control group(s). Rigorous quasi-experimental designs were operationalized as those establishing pre-intervention equivalence between the experimental and control group(s) through the use of multivariate statistical methods or a matched subject research design. For both experimental and quasi-experimental designs, no-treatment control groups did not exclude routine treatment by the criminal justice system. That is, no-treatment could include routine legal interventions such as probation, short jail stay, etc., though it would exclude referral to counseling or alternative programs designed specifically to reduce domestic violence (beyond any deterrent effect of jail or probation).

Second, the intervention must have involved a post-arrest court-mandated intervention that, in part or exclusively, was aimed at the batterer and had as its goal decreasing the batterers' future likelihood of re-assaulting that or other partners. As so defined, pre-trial diversion programs could be included in the study as well. Third, only studies using adult participants of heterosexual intimate domestic violence, whether presently or formerly married, separated, divorced, cohabiting or dating were included in the meta-analysis. Fourth, an outcome measure of repeat domestic violence must have been obtained at least six months post-treatment. The decision to follow offenders for a period post-treatment was based on Dunford's findings that evaluation studies collecting outcome data at the end of treatment were more likely to find effectiveness than those measuring

outcomes for some period post-treatment (Dunford 2000). This suggests that evaluations that are based solely on end-of-treatment assessments should be viewed cautiously. Additionally, the study must have included at least one outcome measure on repeat violence to that or other victims that used something more than offenders' self-reported repeat violence. As such, studies could include victim reports on continued abuse or official measures of recidivism including arrest, charges or convictions. Fifth, the studies need to have been conducted in 1986 or later. Finally, the study needed to have reported sufficient data to permit computation of an effect size.

#### *Search strategy for identification of relevant studies*

Our goal was to identify and include all studies conducted in the United States or elsewhere from 1986 through January 2003 that met our inclusion criteria. Toward this aim, we searched computerized databases and websites (listed below), bibliographies of published reviews of related literature and scrutinized annotated bibliographies of related literature. We conducted searches of the following databases and websites: ERIC, PsycINFO, MEDLINE, Sociological Abstracts, Social Science Index, Social Work Abstracts, Criminal Justice Abstracts, Social Science Citation Index, Lexis Nexis Legal, Lexis Nexis Medical, Dissertation Abstracts International, GPO Monthly Catalog (MOCAT), National Criminal Justice Research Service, Social, Psychological, Criminological and Educational Trials Register (C2-SPECTR), and the PsiTri database of randomized and controlled trials in mental health.

We used 25 keywords in three clusters to search for all experimental and quasi-experimental studies conducted on the effectiveness of court-mandated interventions for domestic violence offenders. Whenever appropriate a 'wildcard' was used so as to search for the root of the word allowing for other possible derivations. (So, for instance, the term 'eval' was used to pick up evaluation, evaluate, evaluating, etc.) Cluster One related to the subject matter. Cluster Two sought to find citations using program keywords. Finally, Cluster Three used keywords related to outcomes. Terms within a cluster were connected with the Boolean 'or' (i.e., an abstract with any one of the terms would be selected) and the clusters were then connected with the Boolean 'and' (i.e., an abstract with at least one of the terms in each cluster would be selected). To make the resulting list more manageable, the search was restricted to titles and abstracts. If the title or abstract looked promising, the entire study was pulled and reviewed. The keywords within each cluster were: (Cluster One) anger management, batter(er/s), domestic assault, domestic violence, family violence, spous(e/al) abuse, physical abuse, Minneapolis Model, Duluth, or intimate partner violence; (Cluster Two) defer(ral/ring/rred), program(s), treatment(s), intervention(s), diversion(ary), or prosecu(te/tion/torial); and (Cluster Three) effect(s/ive/iveness), research(es), outcome(s), eval(uation/uations/ating), experiment(al), quasi(-experimental), random(ly), compar(ison/ing), or match(ed/es/ing). We examined the bibliographies of the following reviews:



Babcock and Taillade (1999), Babcock et al. (2004), Chalk and King (1998), and Davis and Taylor (1999).

The graduate research assistant and the first author reviewed the titles and abstracted those that were identified through the search process. Studies that appeared likely to be eligible were retrieved in their entirety. Where disagreements occurred, the second author was consulted and differences were resolved. The graduate assistant and first author were also responsible for reviewing the full text of all studies retrieved in their entirety to determine final eligibility for the meta-analysis. Again, where there were disagreements or uncertainties regarding the inclusion of a study, the second author's opinion was sought to resolve the eligibility decision.

The above process identified 11,872 titles and abstracts. (Note: These numbers included duplicates.) Fifty-seven studies were retrieved in their entirety for further scrutiny. Fifteen studies representing 10 distinct experimental (four) or quasi-experimental (six) studies were deemed eligible for the meta-analysis (designated with an asterisk in the reference list).

#### *Coding and data management*

Studies determined eligible for inclusion into the systematic review were coded for all relevant data. A four-part coding instrument was used to extract the information (available from the first author). Both authors coded each study and all differences in coding were resolved through negotiation. The coding protocol captured information regarding the nature of the intervention, participant sample, research methods, and outcome results.

To avoid the 'double counting' of findings, two strategies were employed. First, multiple publications of an evaluation were treated as a single study in the synthesis. Second, multiple findings from a single study were categorized by outcome construct and only a single effect per construct was used in any analysis. A decision rule for determining which effect to use in an analysis if multiple effects were available was developed and is explained below in the discussion of results.

#### *Statistical procedures and conventions*

This systematic review used standard meta-analytic methods. More specifically, the odds ratio was used as the effect size for dichotomous outcomes, such as official measures of re-arrest, and the standardized mean difference ( $d$ ) was used for continuous type measures, such as the Conflict Tactics Scales. For ease of presentation, however, the odds ratios were transformed into standardized mean difference type effect sizes. This was done using the methods developed by Hasselblad and Hedges (1995) and involved rescaling the logged odds ratio by a constant. As such, it had no effect on the statistical analyses other than to rescale the values such that they are comparable to the standardized mean difference type effect sizes.

Effects representing unique constructs were analyzed separately. The mean effect size across studies for any given construct was determined by weighting by the inverse variance of the effect size, that is, using the inverse variance weight method. Both fixed and random effects mean effect sizes were computed. The fixed effects model results will only be reported along with the random effects model results if the two models produce substantively different findings (substantially different mean estimate). This is based on an a priori assumption about the distribution of effects across studies that is consistent with the random effects model. These analyses and graphs were performed in Stata with macros written by David B. Wilson that are publicly available <http://mason.gmu.edu/~dwilsonb/ma.html>.

The unit of analysis was the treatment-comparison contrast. Two of the included studies had multiple treatment conditions compared to a single control group (Dunford 2000; Davis et al. 2000). Additionally, Gordon and Moriarty (2003) compared all those mandated to treatment to similar offenders not so mandated and also compared those completing treatment to those who dropped out of treatment. Both of these treatment-comparison contrasts are represented below, although never in the same analysis.

## Results

### *Description of studies*

Four experimental studies and six quasi-experimental studies were identified as meeting the eligibility criteria. The basic research design (i.e., randomized, quasi-experimental with a no treatment comparison group, quasi-experimental with treatment dropouts as the comparison group) and treatment type, number of treatment sessions and weeks, nature of the comparison group, and sample description are reported in Table 1.

All 10 studies were conducted in North America. Nine of these studies were published in peer-reviewed journals, although technical reports were also available for four studies (see reference list). When there was conflicting information between the two sources, data from the non-published technical report was used in the coding of the meta-analysis because that typically provided more detailed information.

All 10 studies evaluated a psychoeducational or cognitive behavioral approach, or some mix of the two approaches targeted at the batterer and delivered in all-male group settings. One study (Dunford 2000) also tested two additional intervention types: a cognitive behavioral group targeted at the male batterer but conducted in conjoint groups as well as a no-program but rigorously monitored intervention. In all but two of the studies (Chen et al. 1989; Dunford 2000) it was noted that the program intervention was accompanied by probation, although in one of these studies (Chen et al. 1989) it seems likely that that was the case as well.

Table 1. Description of studies by author and design type.

| <i>Author by<br/>design type</i>                            | <i>Treatment type</i>                        | <i>Treatment<br/>sessions/weeks</i>        | <i>Comparison<br/>type</i>                                                             | <i>Sample type</i>                                                                       |
|-------------------------------------------------------------|----------------------------------------------|--------------------------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------|
| <b>Randomized</b>                                           |                                              |                                            |                                                                                        |                                                                                          |
| Davis et al. – 8<br>week program                            | Psychoeducational                            | 16/8                                       | Probation and<br>40 hours<br>community<br>service                                      | Convicted<br>batterers – judge,<br>prosecutor and<br>defense must<br>agree to treatment  |
| Davis et al. – 26<br>week program                           | Psychoeducational                            | 26/26                                      | Same as above                                                                          | Same as above                                                                            |
| Dunford – men's<br>group                                    | Cognitive – behavioral                       | 32/52                                      | No treatment                                                                           | Navy sample,<br>incident of<br>domestic violence<br>established,<br>referred to program  |
| Dunford – conjoint                                          | Cognitive – behavioral                       | 32/52                                      | No treatment                                                                           | Same as above                                                                            |
| Dunford – rigorous<br>monitoring                            |                                              | 12/52                                      | No treatment                                                                           | Same as above                                                                            |
| Feder and Forde                                             | Cognitive – behavioral/<br>Psychoeducational | 26/26                                      | Probation                                                                              | All convicted<br>batterers                                                               |
| Palmer et al.                                               | Psychoeducational                            | 10/10                                      | Probation                                                                              | Convicted<br>batterers – unclear<br>how sample drawn                                     |
| <b>Quasi-experimental –<br/>no treatment<br/>comparison</b> |                                              |                                            |                                                                                        |                                                                                          |
| Chen et al.                                                 | Cognitive – behavioral/<br>Psychoeducational | 8 sessions                                 | Non-referred<br>convicted<br>batterers                                                 | Convicted<br>batterers referred<br>to treatment<br>program – unclear<br>how sample drawn |
| Gordon and<br>Moriarty – mandated<br>vs. not                | Psychoeducational                            | 22/22                                      | Probation                                                                              | All convicted<br>batterers                                                               |
| Harrell                                                     | Cognitive – behavioral                       | 10/10                                      | Probation                                                                              | All batterers<br>convicted or<br>given prosecution<br>deferred                           |
| Syers and Edleson                                           | Psychoeducational                            | Batterers not<br>mandated to<br>counseling | All batterers<br>having police<br>contact who<br>could be<br>followed for<br>12 months |                                                                                          |
| <b>Quasi-experimental –<br/>dropouts as comparison</b>      |                                              |                                            |                                                                                        |                                                                                          |
| Dutton                                                      | Cognitive – behavioral                       | 16/16                                      | Treatment<br>dropouts,<br>no-shows<br>and rejects                                      | Convicted<br>batterers – unclear<br>how sample drawn                                     |

Table 1. Continued.

| <i>Author by<br/>design type</i>                    | <i>Treatment type</i>  | <i>Treatment<br/>sessions/weeks</i> | <i>Comparison<br/>type</i> | <i>Sample type</i>                                                                           |
|-----------------------------------------------------|------------------------|-------------------------------------|----------------------------|----------------------------------------------------------------------------------------------|
| Jones and Gondolf                                   | Cognitive – behavioral | 20/20                               | Treatment<br>dropouts      | Batterers in<br>4 treatment<br>programs 79%<br>court-mandated/<br>21% voluntary<br>referrals |
| Gordon and<br>Moriarty – completers<br>vs. dropouts | Psychoeducational      | 22/22                               | Treatment<br>dropouts      | All convicted<br>batterers                                                                   |

Distinct treatment-comparison contrasts within an individual study are listed separately.

The treatment length ranged from a minimum of eight 2-hour sessions (Chen et al. 1989) to a maximum of 32 sessions over the course of a year (Dunford 2000). Treatment length information was not provided by Syers and Edleson (1992). Many of the studies indicated the number of sessions and number of weeks but not the length of the treatment sessions.

The nature of the control group also varied from study to study. The Dunford study (2000) was the most unusual with the control group receiving no intervention whatsoever. Several studies (Feder and Forde 2000;<sup>1</sup> Gordon and Moriarty 2003; Harrell 1991; Palmer et al. 1992) had the control group receiving probation only. The Davis et al. (2000) study used a control group who received 40-hour of community service. The Gordon and Moriarty (2003) study included comparisons to both probation only and treatment no shows and drop-outs. Jones and Gondolf (2002) and Dutton (1986) also used treatment drop-out comparison group designs. Dutton (1986) included men who were rejected from treatment as well as the treatment no-shows and dropouts. Treatment no-shows and drop-outs represented 84% of the sample in Dutton's study and as such is considered a treatment drop-out type study for the analyses below. Finally, one study (Syers and Edleson 1992) did not specify what the control group received beyond not being mandated into counseling.

All but one of the 10 studies used a general civilian population of batterers who were facing or had faced court prosecution for domestic violence. The one exception, Dunford (2000), used men living on a Navy base where an incident of domestic violence had been established and the man had been referred to the program. And all but one study (Jones and Gondolf 2002) used a sample of men who were entirely court-mandated into the batterer program. The Jones and Gondolf (2002) study had a sample that was composed of 79% court-mandated and 21% voluntary clients.

In five studies the representativeness of the sample to the general domestic violence offender population was questionable due to conditions used for inclusion into their sample. In one of the experimental studies (Palmer et al. 1992), inclusion

criteria was suspected of being highly restrictive, in that the resulting sample size was small despite the large jurisdiction from which it was pulled and the long time frame implemented for the study. A second experimental study (Davis et al. 2000) used highly restrictive criteria for inclusion in their sample. In that study, all individuals making up the courtroom workgroup, including the batterer, had to agree to this intervention (versus another non-jail alternative). This, as the researchers noted, led to a pool of more highly motivated offenders than is typically found in the generalized batterer population. In the Dunford study (2000), the men were all living on a naval base with their families and therefore may represent a group with higher stake in conformity than is true of other batterer samples. In one of the quasi-experimental studies (Syers and Edleson 1992) only those men who could be followed 6- and 12-month post-initial police visit were included in the study. This restriction makes it less likely that more marginal batterers would be included in their study. Another quasi-experimental study (Jones and Gondolf 2002) excluded data from one of four sites because the men were deemed at higher risk for subsequent re-offending.

### *Meta-analytic findings*

The effect sizes were analyzed separately by outcome type (official reports and victim reports) and by design type (experimental, quasi-experimental with a no-treatment comparison group, and quasi-experimental with treatment dropouts as the comparison group). Table 2 presents the random-effects mean effect size, 95%

*Table 2.* Random effects mean effect size ( $d$ ) and related statistics for official and victim reported measures of domestic violence by design type.

| Outcome by design type                              | Mean $d$ | 95% CI    |           | $k^a$ | $Q$    |
|-----------------------------------------------------|----------|-----------|-----------|-------|--------|
|                                                     |          | Lower $d$ | Upper $d$ |       |        |
| Official measures                                   |          |           |           |       |        |
| Experiments (randomized)                            | 0.26*    | 0.03      | 0.50      | 7     | 8.19   |
| Quasi-experiments (nonrandomized)                   |          |           |           |       |        |
| No treatment comparison group                       | -0.14    | -0.44     | 0.31      | 4     | 12.00* |
| Treatment dropouts as comparison <sup>b</sup>       | 0.97*    | 0.12      | 1.82      | 3     | 12.00* |
| Victim report measures                              |          |           |           |       |        |
| Experiments (randomized)                            | 0.01     | -0.11     | 0.13      | 6     | 1.84   |
| Quasi-experiments (no treatment comp.) <sup>c</sup> | -0.11    | -0.50     | 0.27      | 1     |        |
| Total                                               | -0.00    | -0.12     | 0.11      | 7     | 2.18   |

\* $p \leq 0.05$ .

<sup>a</sup>Number of effect sizes.

<sup>b</sup>Fixed effects mean effect size was lower (mean  $d = 0.49$ , 95% CI of 0.27–0.71). Although substantially lower in value, this still represents a large effect in this context and leads to the same substantive conclusions.

<sup>c</sup>Fixed effect.

confidence interval, and homogeneity statistic ( $Q$ ) for both outcome types and each design type. The results will be discussed separately for each outcome.

### *Official reports*

Official reports were either official complaints made to the police that may or may not have resulted in an arrest, or actual arrests for domestic violence. If multiple follow-up time points were available, the longest was selected. As can be seen in Table 2, the mean effect size for the experimental (randomized) studies across these 7 comparisons was 0.26. This represents a finding of a moderate reduction in re-offending, with a 95% confidence interval of 0.03–0.50 ( $z = 2.23$ ,  $p = 0.03$ ). Figure 1 indicates a general pattern of positive effects on official reports of repeat victimization in these experimental studies. These estimates varied from a near-zero effect (Davis et al., eight-week program) to large positive effects (Palmer et al., 10-week program; Davis et al., 26-week program). The mean represents a small positive reduction in repeat victimization. This effect roughly represents a reduction in recidivism from 20% to 13%. However, given the small number of studies (four), there is substantial uncertainty regarding the precision of this estimate.

There is additional doubt concerning what the results of one of the studies actually indicates. Specifically, though the Brooklyn Experiment was written as indicating modest support for court-mandated treatment's effectiveness (Davis et al. 2000; Taylor et al. 2001), the findings ran counter to expectations. As noted in their study, treatment completion was higher for the 8-week program than the 26-week program. Yet treatment effects were higher for the men assigned to the 26-week program with an effect size near zero for those assigned to the 8-week program. This differential effect suggests that something other than the batterer program accounted for the positive treatment effect. If the batterer program itself was effective, then the group receiving a higher dose (eight-week program) should have had the better outcome. At the time, Feder and her colleagues speculated that these results were more consistent with a conclusion that supervision, and not

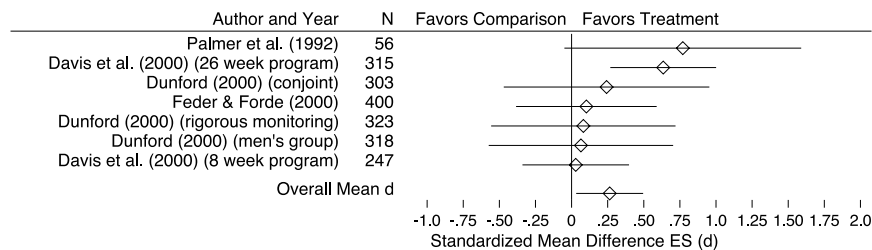


Figure 1. Effect size ( $d$ ) and 95% confidence interval for official measure from experimental (random) studies.

treatment, resulted in the groups' differences in rates of reassaulting (Feder and Forde 2000; Feder and Dugan 2002). In fact, findings from the Brooklyn Experiment have now come to be viewed by the principal investigator as indicative that additional monitoring and not batterer programs were responsible for differences in recidivism between the three groups (Davis, personal communication). Thus, the strongest empirical evidence for the effectiveness of these programs comes from Palmer et al. (1992), a study with a very small sample size (30 men in the batterer program and 26 in the comparison condition). This small sample size leads to a very unstable estimate of the true treatment effect, as is evident in the rather large confidence interval.

We also noticed in coding the experimental studies that the offender population was restricted in some cases, that is, did not reflect the general domestic violence offender population in two studies – the Palmer and Davis studies (see Table 1). Analyzing the official report effect sizes by this distinction shows a lower non-significant overall mean effect size (0.12, with a 95% CI of  $-0.21$  to  $0.44$ ) for the studies using a general domestic violence offender population and a higher mean effect size (0.39, with a 95% CI of  $0.10$ – $0.67$ ) for the studies with a restricted sample. We are unsure what this finding suggests because the specific restrictions placed on one of these samples (Palmer et al. 1992) were not entirely clear. And, as discussed above, the pattern of results between the 8-week and 26-week programs from the Davis et al. (2000) study are not consistent with the hypothesis that batterer intervention programs are effective.

The quasi-experimental studies represented two fundamentally different design types: designs comparing offenders mandated to treatment to those not mandated and designs comparing treatment completers to treatment drop-outs, no-shows and/or rejects. Because the effect that each design is estimating is different, these two design types were analyzed separately. Table 2 indicates that the mean effect size across the former design (not mandated to treatment comparison) was  $-0.14$ , a small negative effect that is statistically not significantly different from zero. As indicated in Figures 2 and 3, these four credible quasi-experimental studies provide a mixed picture (also evidenced by the significant homogeneity test,  $Q$ ), with one study observing a moderate positive benefit, one a small positive benefit and two observing a negative effect of a court-mandated treatment relative to a non-mandated group. These estimates statistically adjust for baseline difference although it is unlikely that all of the important differences between the groups

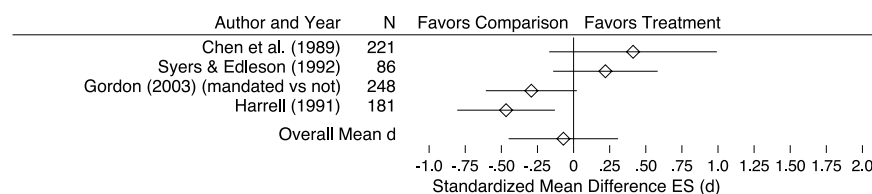


Figure 2. Effect size ( $d$ ) and 95% confidence interval for official measures from quasi-experimental (nonrandomized) studies with a no treatment comparison group.

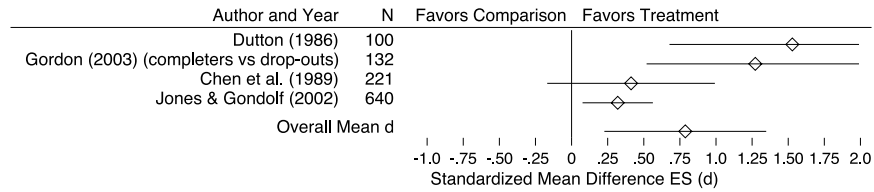


Figure 3. Effect size ( $d$ ) and 95% confidence interval for official measures from quasi-experimental (nonrandomized) studies with treatment dropouts as comparison.

were taken into account. The composite or mean effect has a plausibility range extending from a small negative to a small positive effect.

The second quasi-experimental design type compared batterers who completed a court-mandated treatment program with those who were mandated and were either rejected from treatment, never showed or dropped out. The three studies with this design consistently found a positive and significant effect. That is, abusers mandated to a domestic violence treatment who complete their program re-offend at a substantially lower rate than offenders who were mandated to these programs who did not complete their treatment. Unfortunately, we cannot attribute this difference solely to the impact of treatment, as treatment attendance is likely to be confounded with other important variables. That is, men who attend and complete their treatment may be more highly motivated to change or more fearful of further criminal justice involvement than men who do not complete a treatment program that has been judicially mandated. Differences in rates of recidivism may be attributed, then, to differences existing in the groups prior to the intervention. In other words, the relationship that we think we are observing between treatment non-completion and recidivism may be spurious and due to another, unobserved variable.

#### *Victim reported outcomes*

A concern with official measures is that they may not accurately reflect the amount and severity of ongoing violence. Research consistently indicates that official reports capture only a small fraction of this abuse (Dutton 1988; Straus 1991; Tjaden and Thoennes 2000). As such, the victim is viewed as the best source for information on the offender's continued abuse. Given that, we turn our attention to the seven estimates we have from these studies on the effect of these programs from victim reports of abuse. Three of the four experimental studies measured the victim's reports of their partner's abusive behavior using either the standardized Conflict Tactics Scale (CTS) or the modified Conflict Tactics Scale (CTS2) (Straus et al. 1996). One of the quasi-experimental studies also measured the victim's report of their partner's abusive behavior using a measure similar to



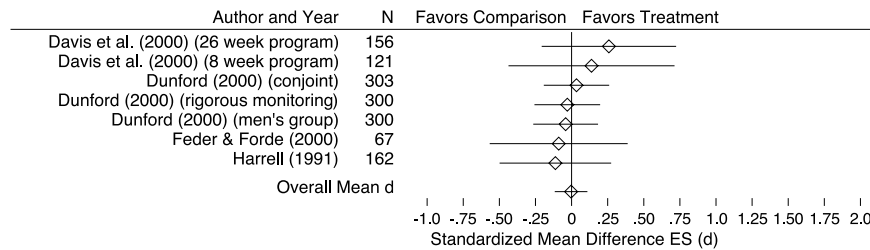


Figure 4. Effect size ( $d$ ) and 95% confidence interval for victim reported measures from experimental and quasi-experimental studies with a no treatment comparison group.

the CTS. For purposes of analysis, we coded all reported subscales and averaged the multiple effect sizes within each treatment-comparison contrast, with the exception of Harrell (1991), where we selected the outcome based on the largest portion of the sample (necessitated by the way in which the researcher reported the data). Thus, the effect size used in Table 2 and Figure 4 represents the mean effect across subscales of the CTS/CTS2 for the comparison of interest. As shown in Table 2, the mean effect size for victim reports in studies using an experimental design was near zero and was not statistically significant. The effect size for quasi-experimental studies showed a small and negative effect for treatment though this finding also was not statistically significant. The distribution of effects is shown in Figure 4. Three of these effects are positive, four are negative, and none are statistically significant. Thus, the outcome measures based on the female intimate partner's report, and the more credible of the quasi-experimental studies, do not replicate the finding of a small but positive benefit of treatment found in the experimental studies using the official measures of re-offending.

## Discussion

This systematic review was based on 10 experimental and quasi-experimental studies. The experimental studies looked at the effect of mandating batterer intervention programs relative to a no-treatment or routine-treatment approach for men facing or convicted of misdemeanor domestic violence charges. Two of the quasi-experimental studies compared men court-mandated to counseling with those not court-mandated (Syers and Edleson 1992; Harrell 1991), two compared court-mandated men who completed treatment to those mandated who did not complete treatment (Dutton 1986; Jones and Gondolf 2002) and one study (Gordon and Moriarty 2003) included both comparisons. All of the evaluated programs used a psychoeducational, feminist oriented and/or cognitive behavioral approach.

The evidence from our meta-analysis is mixed. There is some support for the modest benefits of batterer programs from official reports in the experimental studies, but this effect is smaller if we look only at studies using a general batterer population. Additionally, the effect is absent when victim reported measures are examined. The quasi-experimental studies using a no-treatment comparison also fail to find a positive treatment effect in terms of a reduction in violence when measured with official reports. Finally, quasi-experimental studies using men who were rejected from treatment or who rejected treatment showed a large, positive and significant effect on reducing re-offending. However, we have serious concerns about these latter studies as discussed in detail below.

Our findings differ somewhat from those of Babcock et al. (2004). They concluded, based on their meta-analysis, that these programs have a small but positive effect on abusive behavior. There are several differences between the methods employed in our respective meta-analysis that may account for the differing conclusions. Primarily, Babcock et al. did not separately analyze studies using treatment drop-out designs from other quasi-experimental designs, potentially upwardly biasing the mean effect size for these studies. Both of our results based exclusively on experimental studies are fairly consistent. Babcock et al. reported an effect size of 0.12 when using official reports (fixed effects 95% CI of 0.02–0.22). This is somewhat smaller than our overall mean effect for official reports based on experimental studies but consistent with our estimate from those studies with a representative population. Similarly, Babcock and Steiner indicate a treatment effect of 0.09 (fixed effects 95% CI of –0.02 to 0.21) for victim reported outcomes, slightly higher than our estimate (0.01) but neither estimate is statistically significant.

We have four main concerns regarding the findings. First, we question the generalizability of these findings to general convicted batterer populations. Second, we believe there is a potential bias inherent in using official records to measure continued abuse. Third, the victim reports suffer from low reporting rates in these studies, raising concerns about the validity of these estimates. Finally, we question the validity of the quasi-experimental studies that compare treatment completers to rejects, no-shows and dropouts. Each of these concerns is addressed below.

We judged two studies (Davis et al. 2000; Palmer et al. 1992) as having samples that were restricted in a manner that reduced the representativeness of their findings to a general batterer population. Studies that did not have restrictions limiting who was included in the batterer program probably better represented the ‘typical’ convicted batterer. Our analysis indicates that these latter studies had a lower overall mean effect size for official reports of domestic violence than the studies using a restricted sample. Importantly, the mean effect for the more representative studies was not statistically significant, raising the possibility that the overall positive findings of Figure 1 was in part a function of a restricted (possibly more motivated or perhaps ‘creamed’) sample of batterers. This may indicate that batterer intervention programs work for a selected (presumably more motivated) subset of offenders. The evidence on this issue is weak for two reasons:

(1) we do not actually know the motivation levels of the men in the different studies, and (2) the Davis et al. study had inconsistent results across two similarly motivated groups receiving the same intervention, differing only in the number of weeks over which the program was spread. Thus, we believe that there is insufficient data for any strong conclusion on this issue.

The heavy reliance on official measures in all of these studies is also highly problematic. Official measures are dependent on a victim's willingness to file a complaint or call the police. This raises the possibility that assignment to court-mandated treatment versus a no-treatment control group may differentially affect the victim's willingness to contact criminal justice officials when future abuse occurs. (What Cook and Campbell (1979) refer to as an instrumentation effect.) A victim may not report her partner's abuse for a number of reasons. This includes the possibility that she might prefer to see her partner continue in treatment where she believes it will eventually lead to changes in his abusive behavior rather than take the risk of reporting his continued abuse and see him go to jail. Alternately, a victim may resent the criminal justice system's intrusion into her life in the form of mandating a treatment that she is then responsible to pay for. Most programs require the abuser to pay for the treatment and by extension that means that it is the family that pays for the treatment (Zorza 2003). If the treatment is viewed by a victim as ineffective, it may make her critical and suspicious of the system and less likely to cooperate in the case of reporting future incidences of abuse. We have no empirical evidence that this occurs but the dependence of official reports on the behavior of the victim allows for the possibility that the different rates noted between batterers in the treatment and comparison conditions may reflect a measurement artifact and not a genuine treatment effect. This possibility is strengthened by the different findings obtained in these studies depending upon whether official reports or victim reports are used as the outcome measure.

The high rate of victim attrition in many of these studies is another concern. The victim is usually viewed as the best source for information on the offender's continued abuse. Victim reports of abuse via standardized measures such as the Conflict Tactics Scale are less likely to be affected by the issues raised regarding official reports of continued abuse, provided that the victim is convinced of the confidential nature of her responses. Unfortunately, the percentage of victims responding to follow-up surveys in these studies is low, seriously undermining their utility in establishing the effectiveness (or ineffectiveness) of these programs.

The attrition for victim report for the effect sizes shown in Figure 4 was roughly 30% for the Dunford (2000), roughly 50% for the Davis et al. (2000), roughly 80% for the Feder and Dugan (2002) and 59% for Harrell (1991). High attrition raises the possibility that the victims lost to follow-up in the treatment group may differ in meaningful ways from those lost to follow-up in the control group. Thus, the absence of an effect for the victim report measures may reflect that the programs are truly ineffective or, alternately, that there is a positive or negative effect that is masked by differential attrition.

The problem of high rates of victim attrition becomes critical in light of research indicating that certain victims of domestic violence are more likely to be

lost in the research follow-up than are others. This research strongly suggests that women victims of domestic violence who are more difficult to retain in follow-up research are both more marginal and more likely to be more frequently and severely abused (Sullivan et al. 1996). There is also research that indicates that men who are more marginal are both less likely to obey a court-mandate to treatment and more likely to continue to abuse their partners (Feder and Dugan 2002). If we can assume that more marginal women are more likely to be partnered with more marginal men, then the need for maintaining contact with a high percentage of victims when assessing the effectiveness of these spouse abuse abatement programs becomes even more apparent. This may be important to the extent that some research has indicated that factors associated with the abuser's stake in conformity is associated with the likelihood that an intervention will be successful in reducing subsequent violence (Berk et al. 1992; Sherman 1992). At best, this attrition reduces the generalizability of the findings from victim reported outcomes to a subset of the domestic violence offender population. At worst, there may be differential loss of these marginal women from the treatment and control groups, producing bias in the findings.

Finally, we note the difficulty with using treatment dropouts as a control group, even once statistical controls have been introduced. Two specific problems occur with this type of study design, one with the construct of what is being evaluated or tested and the other with the adequacy of the statistical models in adjusting for initial group differences. First, these studies are trying to estimate the affect of full participation in the batterer intervention program above and beyond the court mandate. In other words, they look to answer the question, "Among men who are court-mandated to batterer intervention, do those who choose to attend and complete this program do better than those who do not?" Although this may be of interest to program providers and developers, it does not address the broader issue of the likely reduction in domestic violence as a function of a policy to mandate such treatment. That is, "whether court-mandated batterer intervention programs reduce offenders' likelihood of re-offending." Addressing the latter question is critical to knowing whether court-mandated domestic violence interventions are beneficial to society.

Second, that these studies produce treatment effect estimates that are large given the population and nature of the problem clearly establishes that men who complete these programs recidivate at a lower rate than men who do not. The question is what to make of this empirical finding. The statistical models employed by these studies attempt to adjust for selection differences between the groups of men. To produce unbiased estimates, however, these models need to fully account for the selection process, that is, the reasons why some men attended treatment and others did not. We do not believe that these equations adequately model the selection process. Potentially important variables, such as motivation for treatment, were not included. The positive treatment effect estimate may indicate that the treatment is effective for motivated offenders, though we cannot conclude this since we do not have comparisons in any of these quasi-experimental studies with motivated offenders who were not mandated and did not receive treatment. Equally

plausible, these findings may simply reflect that the subset of offenders who will complete mandated treatments are less likely to re-offend, with or without the treatment (i.e., these programs may have ‘creamed’ those offenders who are least likely to re-offend regardless of what action is taken).

### **Conclusion**

The findings from this meta-analysis combined with the caveats above raise questions as to the value of these programs. While additional research is needed, the meta-analysis does not offer strong support that court-mandated treatment to misdemeanor domestic violence offenders reduces the likelihood of further reassault.

### *Research implications*

The research implication growing out of this synthesis is that additional experiments need to be conducted to more clearly decipher the effectiveness of court-mandated batterer intervention programs. If we are to test the ability of courts to mandate change, these future experiments must ensure samples of batterers that are representative of the larger convicted batterer population rather than a smaller subset of selected batterers. Additionally, these studies must attend to the importance of maintaining high victim retention, so as to better ascertain any positive or negative effects from this mandated intervention. Finally, additional research is needed to better understand the validity and reliability of official report and victim report measures used in these studies and how they might be affected by treatment assignment.

### *Policy implications*

Intervening in the lives of others is a risky business, particularly when the individuals participating in the social intervention are mandated by a court of law to do so. As such, it is incumbent upon us to ensure that we are not inadvertently making things worse for those we are seeking to help. At this point, the existing evidence cannot ensure that these programs are, in fact, helpful and not harmful.

There is no doubt that “There is a tremendous sense of urgency and alarm in the treatment of domestic violence – and rightly so. After all, protecting the physical and emotional safety of women and their children is the first priority. Consequently, clinicians feel a primary obligation to ‘do something immediately and decisively to halt and prevent violence’” (Jennings 1987: 204). But as the above review has indicated, doing something may not help. As McCord so wisely noted, “Unless social programs are evaluated for potential harm as well as benefit,

safety as well as efficacy, the choice of which social programs to use will remain a dangerous guess" (McCord 2003: 16).

It is clear that we need to be guided by rigorous research in helping us set our course. While better research is needed to determine the effectiveness of court-mandated batterer intervention programs, the results from the meta-analysis do not provide confidence that these programs will be found to be effective. Therefore, it would prove beneficial for the criminal justice system to begin looking at other types of interventions for addressing the problem of domestic violence. However, these interventions must be tied to rigorous evaluations to determine their full impact. In other words, we recommend the use of pilot studies joined to an experimental design, as was suggested almost 20 years ago by Berk et al. (1985), as the preferred path for finding effective programs that can effectively meet the challenge that intimate partner violence presents. Such a course would be especially prudent in these times of limited resources. More than that, victims and taxpayers deserve such evidence-based decision making.

Unfortunately, what we are suggesting is not possible in many jurisdictions today in that their statutes require that, upon conviction for domestic violence, individuals must be mandated into a batterer intervention program, not atypically based upon the Duluth Model (Babcock and Taillade 2000). The end result is that judges, prosecutors and probation officers continue to send batterers to these treatment programs, even if they have grave doubts about their effectiveness. Alternate programs cannot be implemented and tested even as evidence builds indicating that batterer intervention programs, at least as designed and implemented today, may not be effective.

### Acknowledgements

We thank Sabrina Austin for her work searching, retrieving, and reviewing for eligibility the studies used in this meta-analysis.

This work was supported in part by funding from the Smith Richardson foundation.

### Note

- 1 The first listed author on this review (Lynette Feder) was the primary investigator of the Broward Experiment assessing the effectiveness of batterer intervention programs in South Florida. To best counter the potential conflict of interest, the review was done as transparently as possible. Additionally, the researcher chose to collaborate with a colleague to ensure against any bias. Finally, it was decided before beginning the project that where there were disagreements between the first and second authors, an expert arbitrator would be brought in to resolve any conflicts.

## References

- Adams, D. & McCormick, A. (1982). Men unlearning violence: A group approach based on the collective model. In M. Roy (Ed.), *The abusive partner: An analysis of domestic battering* (pp. 170–197). New York: Van Nostrand-Reinhold.
- Babcock, J. C. & Taillade, J. (2000). Evaluating interventions for men who batter. In J. Vincent & E. Jouriles (Eds.), *Domestic violence: Guidelines for research-informed practice* (pp. 37–77). Philadelphia: Jessica Kingsley.
- Babcock, J. C., Green, C. E. & Robie, C. (2004). Does batterers' treatment work? A meta-analytic review of domestic violence treatment. *Clinical Psychology Review* 23(8), 1023–1053.
- Berk, R., Boruch, T., Chambers, F., Rossi, P. & Witte, S. (1985). Social policy experimentation: A position paper. *Evaluation Review* 9(4), 387–429.
- Berk, R., Campbell, A., Klap, R. & Western, B. (1992). The deterrent effect of arrest in incidents of domestic violence: A Bayesian analysis of four field experiments. *American Sociological Review* 57(5), 698–708.
- Brisson, N. (1981). Battering husbands: A survey of abusive men. *Victimology* 6, 338–344.
- Chalk, R. & King, P. (1998). *Violence in families: Assessing prevention and treatment programs*. Washington, DC: National Academy Press.
- \*Chen, H., Bersani, C., Myers, S. & Denton, R. (1989). Evaluating the effectiveness of a court sponsored treatment program. *Journal of Family Violence* 4, 309–322.
- Cook, T. D. & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.
- Cromwell, N. & Burgess, A. (1996). *Understanding violence against women*. Washington DC: National Academy Press.
- Davis, R. & Taylor, B. (1999). Does batterer treatment reduce violence? *Women and Criminal Justice* 10, 69–93.
- \*Davis, R. C., Taylor, B. G. & Maxwell, C. D. (2000). *Does batterer treatment reduce violence? A randomized experiment in Brooklyn*. Washington, DC: National Institute of Justice.
- \*Dunford, F. W. (2000). The San Diego Navy experiment: An assessment of interventions for men who assault their wives. *Journal of Consulting and Clinical Psychology* 68, 468–476.
- \*Dutton, D. (1984). Interventions into the problem of wife assault: Therapeutic, policy and research implications. *Canadian Journal of Behavioral Science* 16(4), 281–297.
- Dutton, D. (1986). Wife assaulter's explanations for assault: The neutralization of self-punishment. *Canadian Journal of Behavioral Science* 18(4), 381–390.
- Dutton, D. (1988). Research advances in the study of wife assault: Etiology and prevention. *Law and Mental Health* 4, 161–220.
- Dutton, D. & McGregor, B. (1991). The symbiosis of arrest and treatment for wife assault: The case for combined intervention. In M. Steinman (Ed.), *Woman battering: Policy responses* (pp. 131–154). Cincinnati, OH: Anderson Publishing Company.
- Eisikovits, Z. & Edleson, J. (1989). Intervening with men who batter: A critical review of the literature. *Social Service Review* 63, 384–414.
- Feder, L. (1997). Domestic violence and police response in a pro-arrest jurisdiction. *Women and Criminal Justice* 8(4), 79–98.
- \*Feder, L. & Dugan, L. (2002). A test of the efficacy of court mandated counseling for domestic violence offenders: The Broward Experiment. *Justice Quarterly* 19(2), 343–375.



- Feder, L. & Forde, D. (2000). *A test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward Experiment* (Final report, Grant NIJ-96-WT-NX-0008). Washington, DC: National Institute of Justice.
- Farley, D. & Magill, J. (1988). An evaluation of a group program for men who batter. *Social Work With Groups* 11(3), 53–65.
- Feazell, C., Mayers, R. & Deschner, J. (1984). Services for men who batter: Implications for programs and policies. *Family Relations* 33, 217–223.
- Ford, D. & Regoli, M. J. (1993). The criminal prosecution of wife assaulters. In Z. Hilton (Ed.), *Legal responses to wife assault: Current trends and evaluation* (pp. 127–164). Newbury Park, CA: Sage.
- Gondolf, E. (1987). Evaluating programs for men who batter: Problems and prospects. *Journal of Family Violence* 2(1), 95–108.
- Goolkasian, G. (1986). *Confronting domestic violence: The role of criminal court judges*. Washington, DC: National Institute of Justice.
- \*Gordon, J. A. & Moriarty, L. J. (2003). The effects of domestic violence batterer treatment on domestic violence batterer treatment and domestic violence recidivism. *Criminal Justice and Behavior* 30(1), 118–134.
- Hamberger, L. K. & Hastings, J. (1989). Counseling male spouse abusers: Characteristics of treatment completers and dropouts. *Violence and Victims* 4(1), 275–286.
- Hamberger, L. K. & Hastings, J. (1993). Court-mandated treatment of men who assault their partner. In Z. Hilton (Ed.), *Legal responses to wife assault: Current trends and evaluation* (pp. 188–229). Newbury Park, CA: Sage.
- \*Harrell, A. (1991). *Evaluation of court-ordered treatment for domestic violence offenders* (Final report). Washington, DC: National Institute of Justice.
- Hasselblad, V. & Hedges, L. V. (1995). Meta-analysis of screening and diagnostic tests. *Psychological Bulletin* 117, 167–178.
- Healey, K. & Smith, C. (1998). *Batterer programs: What criminal justice agencies need to know*. Washington, DC: National Institute of Justice.
- Healey, K., Smith, C. & O'Sullivan, C. (1998). *Batterer intervention: Program approaches and criminal justice strategies*. Washington, DC: Department of Justice.
- Hilberman, E. (1980). Overview: The “wife-beater’s wife” reconsidered. *American Journal of Psychiatry* 137(11), 1336–1347.
- Hirschel, J. D. & Hutchinson, I. (1992). Female spouse abuse and the police response: The Charlotte, North Carolina Experiment. *Journal of Criminal Law and Criminology* 83(1), 73–119.
- Hotaling, G. & Sugarman, D. (1986). An analysis of risk markers in husband to wife violence: The current state of knowledge. *Violence and Victims* 1(2), 101–124.
- Jennings, J. (1987). History and issues in the treatment of battering men: A case for unstructured group therapy. *Journal of Family Violence* 2(3), 193–213.
- Johnson, J. & Kanzler, D. (1993). Treating domestic violence: Evaluating the effectiveness of a domestic violence diversion program. *Studies in Symbolic Interaction* 15, 271–289.
- \*Jones, A. S. & Gondolf, E. W. (2002). Assessing the effect of batterer program completion on reassault: An instrumental variables analysis. *Journal of Quantitative Criminology* 18(1), 71–98.
- Langan, P. & Innes, C. (1986). *Preventing domestic violence against women*. Washington, DC: National Institute of Justice.
- McCord, J. (2003). Cures that harm: Unanticipated outcomes of crime prevention programs. *Annals of the American Academy of Political and Social Science* 587, 16–30.
- Miller, T., Cohen, M. & Wiersema, B. (1996). *Victim costs and consequences: A new look*. Washington, DC: National Institute of Justice.



- \*Palmer, S., Brown, R. & Barrera, M. (1992). Group treatment program for abusive husbands: Long-term evaluation. *American Journal of Orthopsychiatry* 62(2), 276–283.
- Pence, E. (1983). The Duluth Domestic Abuse Intervention Project. *Hamline Law Review* 6, 247–275.
- Pirog-Good, M. & Stets-Kealey, J. (1985). Male batterers and battering prevention programs: A national survey. *Response* 8, 8–12.
- Rennison, C. R. & Welchans, S. (2000). *Intimate partner violence*. Washington, DC: National Institute of Justice.
- Roberts, A. (1982). A national survey of services for batterers. In Maria Roy (Ed.), *The abusive partner: An analysis of domestic battering* (pp. 230–243). New York: Van Nostrand-Reinhold.
- Rosenfeld, B. (1992). Court-ordered treatment of spouse abuse. *Clinical Psychology Review* 12, 205–226.
- Saunders, D. (1996). Interventions for men who batter: Do we know what works? *In Session: Psychotherapy in Practice* 2(3), 81–93.
- Sherman, L. (1992). The influence of criminology on criminal law: Evaluating arrests for misdemeanor domestic violence. *Journal of Criminal Law and Criminology* 83, 1–45.
- Snyder, D. & Scheer, N. (1981). Predicting disposition following brief residence at a shelter for battered women. *American Journal of Community Psychology* 9, 559–566.
- Sonkin, D. J. (1988). The male batterer: Clinical and research issues. *Violence and Victims* 3(1), 65–79.
- Straus, M. (1991). Conceptualization and measurement of battering: Implications for public policy. In M. Steinman (Ed.), *Woman battering: Policy responses* (pp. 19–47). Cincinnati, OH: Anderson.
- Straus, M., Hamby, S., Boney-McCoy, S. & Sugarman, D. (1996). The revised Conflict Tactics Scale (CTS2): Development and preliminary psychometric data. *Journal of Family Issues* 17(3), 283–316.
- Sullivan, C., Rumptz, M., Campbell, R., Eby, K. & Davidson, W. (1996). Retaining participants in longitudinal community research: A comprehensive protocol. *Journal of Applied Behavioral Science* 32(3), 262–276.
- \*Syers, M. & Edleson, J. (1992). The combined effects of coordinated criminal justice intervention in woman abuse. *Journal of Interpersonal Violence* 7, 490–502.
- Taylor, B., Davis, R. & Maxwell, C. (2001). The effects of a group batterer treatment program: A randomized experiment in Brooklyn. *Justice Quarterly* 18(1), 171–201.
- Tjaden, P. & Thoennes, N. (2000). Prevalence and consequences of male-to-female and female-to-male intimate partner violence as measured by the National Violence Against Women Survey. *Violence Against Women* 6(2), 142–161.
- Tolman, R. & Bennett, L. (1990). A review of quantitative research on men who batter. *Journal of Interpersonal Violence* 5, 87–118.
- Tolman, R. & Edleson, J. (1995). Intervention for men who batter: A review of research. In S. Stith & M. Straus (Eds.), *Understanding partner violence: Prevalence, causes, consequences and solutions* (pp. 262–273). Minneapolis, MN: National Council on Family Relations.
- Weisburd, D., Lum, C. & Petrosino, A. (2001). Does research design affect study outcomes in criminal justice? *Annals of the American Academy of Political and Social Science* 578, 50–70.
- Widom, C. S. (1992). *The cycle of violence*. Washington, DC: US Department of Justice.
- Zorza, J. (2003). New research: Broward County Experiment shows no benefit from batterer intervention programs. *Domestic Violence Report* 8, 23–25.

**About the authors**

**Lynette Feder**, Ph.D. is an Associate Professor in the Criminal Justice and Criminology division at Portland State University. Her research interests are directed at rigorously evaluating criminal justice interventions to address both policy questions as well as underlying theoretical issues in areas including juvenile delinquency, discretion and discrimination in the criminal justice system, and domestic violence.

**David B. Wilson**, Ph.D., is an Assistant Professor in the Administration of Justice Program at George Mason University. His research interests are the effectiveness of offender rehabilitation and crime prevention efforts, program evaluation methodology, quantitative methods, and meta-analysis.



# BATTERER INTERVENTION SYSTEMS IN CALIFORNIA

## AN EVALUATION



ADMINISTRATIVE OFFICE  
OF THE COURTS

---

OFFICE OF COURT RESEARCH





# BATTERER INTERVENTION SYSTEMS IN CALIFORNIA AN EVALUATION

---

Dag Macleod

Ron Pi

David Smith

Leah Rose-Goodwin



ADMINISTRATIVE OFFICE  
OF THE COURTS

---

OFFICE OF COURT RESEARCH

Judicial Council of California  
Administrative Office of the Courts  
Office of Court Research  
455 Golden Gate Avenue  
San Francisco, California 94102-3688

This project was supported by Award No. 2005WGBX0004 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice and with the support of the Administrative Office of the Courts, Executive Office Programs Division, Office of Court Research. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the Department of Justice or the Administrative Office of the Courts.

Copyright © 2009 by Judicial Council of California/Administrative Office of the Courts. All rights reserved.

Except as permitted under the Copyright Act of 1976 and as otherwise expressly provided herein, no part of this publication may be reproduced in any form or by any means, electronic or mechanical, including the use of information storage and retrieval systems, without permission in writing from the copyright holder. Permission is hereby granted to nonprofit institutions to reproduce and distribute this publication for education purposes if the copies credit the copyright holder.



## Acknowledgements

The authors of this report would like to express their deep gratitude to the judges, court staff, department of probation staff, and batterer intervention program directors and staff in Los Angeles, Riverside, San Joaquin, Santa Clara and Solano Counties. The research that informs this report would not have been possible without their generous support. These professionals have dedicated their lives to eradicating domestic violence. Despite their busy schedules and other commitments, they made the time to actively assist with the collection (and sometimes the interpretation) of data while providing constructive criticism and challenging the research team to think harder about the complex issues involved in the justice system's intervention in domestic violence cases.

The authors would also like to express their appreciation to their colleagues at the Administrative Office of the Courts who contributed directly and indirectly to the research design, data collection, and analysis. Current and former Office of Court Research staff Yueh-Wen Chang, Selina Hung, Kevin O'Connell and Ruth White all assisted with data collection at various points in the project. Members and staff of the Domestic Violence Practice and Procedure Task Force were also instrumental in assisting with the project. External support was provided by Professor Eric Mankowski, who provided the team with a psycho-social measurement instrument to assess changes in offender attitudes and beliefs. The Research Advisory Committee, consisting of Professors Donald G. Dutton, Theodore Eisenberg, and Lorena Garcia, and Andrew R. Morral of the RAND Corporation, offered invaluable critique on the design and analysis of the research. Any errors or omissions remain the sole responsibility of the authors.

Finally, the authors wish to thank Bernard Auchter, grant monitor at the Department of Justice, who provided essential logistical support in navigating the grant management process as well as thoughtful comments related to the original research design.



## Authors

Dag MacLeod, Ph.D., is the manager of the Office of Court Research, Executive Office Programs Division of the Administrative Office of the Courts (AOC) of California and has worked for the AOC since 1999. He received his Ph.D. in sociology from the Johns Hopkins University.

Ron Pi, Ph.D., is a Supervising Research Analyst in the Office of Court Research and has worked for the AOC since 2000. He received his graduate degree in political economy from the University of Texas at Dallas and bachelor's degree in foreign languages and literature from the National Taiwan University.

David A. Smith, Ph.D., is a Senior Research Analyst with the Office of Court Research and has worked for the AOC since 2002. He received his Ph.D. from the University of Michigan in psychology and a master's from the University of California, Berkeley in an interdisciplinary program in public health (social epidemiology) and environmental planning.

Leah Rose-Goodwin, M.P.P., has worked for the Office of Court Research since 2004 and is currently a Senior Research Analyst. She received her Master's Degree in public policy from the University of California, Berkeley and a Bachelor's Degree in political science from the University of California, San Diego.

# Table of Contents

|                                                                     |     |
|---------------------------------------------------------------------|-----|
| Acknowledgments                                                     | i   |
| Authors                                                             | ii  |
| Executive Summary                                                   | v   |
| Chapter 1. Evaluation of California’s Batterer Intervention Systems | 1   |
| Chapter 2. Five Batterer Intervention Systems in California         | 17  |
| Chapter 3. Batterer Intervention Program Content                    | 33  |
| Chapter 4. Offender Profiles                                        | 51  |
| Chapter 5. Analysis of Systems Impacts                              | 67  |
| Chapter 6. Policy Issues and Research Implications                  | 97  |
| Appendices                                                          | 107 |



# **Batterer Intervention Systems in California:**

## **Executive Summary**

Domestic violence represents both a serious criminal justice and public health problem. Every year in California over 100,000 arrests are made for misdemeanor and felony domestic violence charges while countless additional cases of intimate-partner violence go unreported. The social, economic, and personal costs of domestic violence make it a critical area for evaluating the effectiveness of the justice system response to this crime.

Since 1994, California law has required defendants who are convicted and granted probation in domestic violence cases to complete a certified batterer intervention program (BIP). In addition, recognizing the severity of the problem of intimate-partner violence and the unique challenges these cases present, many superior courts in California have adopted specialized procedures for handling domestic violence cases such as using dedicated calendars and holding periodic review hearings with offenders.

This study seeks to take advantage of the fact that each jurisdiction in California manages its domestic violence caseload somewhat differently. We begin by documenting the different ways that courts, departments of probation, and BIPs intervene with domestic violence offenders in a sample of five jurisdictions—Los Angeles, Riverside, San Joaquin, Santa Clara and Solano. We then compare the efficacy of the justice system response across jurisdictions, looking primarily at differences in rates of program completion and re-offense by offenders.

Drawing on a sample of over a thousand men enrolled in treatment programs across the five jurisdictions, this study is the largest of its kind ever conducted.<sup>1</sup> It lays the foundation for improving the justice system response to domestic violence and for future research to untangle the complex relationships among the individual characteristics of men who commit domestic violence, the BIPs that are charged with treating these men, and the efforts of courts and departments of probation to hold offenders accountable and ensure victim safety.

## ***Findings***

- The men who find their way into the justice system and ultimately enroll in BIPs appear to be non-representative of the larger social problem of domestic violence. The sample of men convicted of domestic violence offenses drawn for this study generally had low levels of educational attainment, were poor, majority Hispanic, and had lengthy criminal records;
- Slightly more than one third of the men convicted of domestic violence in our sample report that they still live with their victim; about one third of the men reported that they live with children;

- BIPs appear to incorporate multiple approaches to intervention with domestic violence offenders into their programs, integrating components of cognitive behavioral therapy, the Duluth model and other methods that they determine are appropriate and effective;
- The educational topics that BIPs identified as important to helping offenders end their abuse appear to be consistent with the legislative requirements for these programs;
- Offenders' rates of program completion varied across different BIPs. The reason for this, however, appears to be in part that the characteristics of men who are enrolled in different BIPs varies systematically across programs. The statistical significance of the differences in program completion across BIPs declines as additional, individual-level variables are added to the model;
- In contrast to the weak correlation between program completion and BIP, there is no statistical association at all between programs and an offender's likelihood of re-offense;
- For offenders who successfully completed the 52-week BIP, attitudes and beliefs showed small, positive, changes along a number of dimensions including taking greater personal responsibility, understanding the effect of abuse on others, and anger management;
- The strongest predictors of whether or not men were re-arrested following intake in a BIP were individual characteristics of the offenders, not the characteristics of jurisdictions or BIPs in which offenders were enrolled.<sup>2</sup> Men who were more educated, older, had shorter criminal histories, and did not display clear signs of drug or alcohol dependence had a lower likelihood of re-arrest;
- Whether probation or the court is primarily responsible for oversight of the offenders made no difference in the likelihood of re-arrest. This finding is similar to the conclusion of a recent study in which judicial supervision of domestic violence offenders—with comparisons between supervision of different intensity and compared with no supervision at all—was found not to make any significant difference on recidivism 12 months after sentencing;<sup>2</sup>
- Even after controlling for individual characteristics, two jurisdictions showed statistically significant differences in outcomes for offenders. Using Los Angeles as the base for comparison, offenders in Solano County had a likelihood of re-arrest at 12 months after intake that is one-third the likelihood of offenders in Los Angeles County, while offenders in Santa Clara County were 1.6 times as likely to be arrested as offenders in Los Angeles.

## ***Implications for Policy***

- Because of the salience of individual characteristics in predicting program completion and re-offense, enhanced risk and needs assessment at intake may improve offender treatment.

Penal Code §1203.097(b)(1) lays out detailed offender assessment requirements but limits these to offenders who are on formal probation. The collection of information on basic risk and needs factors of offenders who are informally as well as those that are formally supervised by probation would allow BIPs to tailor their treatment more narrowly.

- Drug/alcohol treatment may be essential to help offenders end their abuse.

Indicators of risk for drug and alcohol abuse are strong predictors of non-completion of batterer intervention programs and senior program staff in the BIPs generally agreed that addressing the topic of alcohol and drug abuse is important to helping offenders end their domestic violence. Because many BIPs have limited resources and little leverage over offenders enrolled in their programs, it may be useful for departments of probation and the courts to consider how best to support BIPs in requiring batterers at risk for substance abuse to attend some reasonable form of drug/alcohol treatment in conjunction with their enrollment in the BIP.

- The current BIP fee structure may hinder differentiated case management.

Enhanced risk and needs assessment combined with heightened attention to drug/alcohol abuse suggest that the justice system may need to engage in more differentiated case management with domestic violence offenders. One more piece of the puzzle of differentiated case management has to do with fees.

The current method of assessing and paying fees, all managed at the BIP level, may pose a barrier to a differentiated treatment model because Pen. Code §1203.097 mandates probation departments to evenly allocate referrals of indigent clients among approved programs. Thus, the effort to assign the right socioeconomic balance to different programs may undermine efforts to assign men to programs on the basis of the characteristics that put them most at risk for re-offense.

Moreover, it is not clear that enough higher-income men are available in the system to cross-subsidize the costs of the lower-income men in programs. Creating a more differentiated treatment model might require an exploration of alternative fee distribution and payment plans. This might grant BIPs the financial freedom to accept enrollments on the basis of service need rather than have to consider a client's ability to pay.

## ***Implications for Research***

- The effort to understand the impact of the justice system as a whole is hampered by variation *within* jurisdictions.

Differences in court practice from location to location within jurisdictions, as well as large variability in outcomes across BIPs within jurisdictions undercut efforts to evaluate the justice system response. Instead, in some cases we have findings related to different systems within a single jurisdiction that cannot be completely disentangled.

Further integration of the qualitative data will assist with the interpretation of the findings. Once the qualitative differences within jurisdictions are better understood, quantitative analysis that excludes outlying court locations where these introduce too much variability might be a fruitful path for recapturing the system perspective that motivated this study. Given the clustering of large numbers of offenders in specific courts and in some specific BIPs, this may be a near- to medium-term follow-up study with this data set.

- Clearer specification of system intervention measures is needed.

System intervention measures such as “probation contact,” “court review,” or even “attendance” at a BIP are all inherently limited by the variability in how these interventions occur across locations. Consistent with the other observations here, more qualitative information on what these variables really are in practice—whether probation contact is a face-to-face interview at the department of probation as opposed to a check-in by telephone or whether the review at the trial court is in open court in front of a judge or handled by a courtroom clerk—will assist in distinguishing among different systems.

- More information on BIPs is needed to understand and identify promising practices.

The challenge of interpreting outcomes given the variability across jurisdictions is compounded by variability across BIPs. Although this study captured measures of BIP priorities for teaching and training related to different elements of the intervention, the findings did not show sufficient variability to introduce the data into our quantitative models and to begin teasing out the effects that these programs produce on offender outcomes.

In the future, this information will need to be combined with independent measures if we are to clearly understand the approach intervention programs are taking in their work with clients. Further, we need to learn more about BIPs as practitioner groups and/or organizations in terms of their staffing levels and role differentiation, the training and professional experience levels of program staff, the supplementary services BIPs are able to provide clients directly or indirectly, and the resources these organizations have at their disposal to sustain their work with batterers. Such information is essential to our ability to understand BIPs in their various organizational forms, as well as to identify promising program approaches and practices.

## ***Layout of the Report***

This report is organized to isolate and describe the variation that is found at different levels of analysis in the five study jurisdictions. After introducing the study in Chapter 1, Chapters 2 through 4 move from the highest level of analysis – the jurisdictional differences across counties – to successively lower levels of analysis – the BIPs, and then the individuals within the programs. In Chapter 5, the variables that are described in the preceding chapters are brought together for the final evaluation of outcomes.

- Chapter 1 outlines the research design and methodology employed. This chapter places the study in the context of previous research on this topic, lays out the logic model for the study, defines the study population, and discusses the various types and sources of data collected;
- Chapter 2 provides an overview of the five jurisdictions in the study including both qualitative description of the coordination of domestic violence cases among justice system partners and quantitative measures of court and probation oversight of offenders;
- Chapter 3 looks more closely at the actual content of BIP curriculum and teaching strategies in the study jurisdictions. This chapter describes the findings of a survey of 45 BIPs in the study jurisdictions on the educational topics and teaching methods employed by BIPs;
- Chapter 4 describes the offenders enrolled in the study, including detailed information on age, race/ethnicity, family living arrangements, educational attainment, income levels, criminal history, and risk of drug/alcohol dependence;
- Chapter 5 brings together all of the variables described in the preceding chapters to evaluate the impact of the jurisdiction on two primary outcome measures: program completion and re-arrest. This chapter also evaluates changes in attitudes and beliefs among a smaller sub-sample of men who completed the BIP during the study period;
- Chapter 6 summarizes the study findings and looks at the implications for both policy and research.



## Endnotes Executive Summary

---

1. This study looks exclusively at men who committed domestic violence offenses against female partners in an effort to understand the justice system response to the largest proportion of the domestic violence caseload and to minimize the variability within the sample.
2. All findings discussed in this Executive Summary are statistically significant at a level of .01 or .05 unless otherwise noted.
3. Melissa Labriola, Michael Rempel, and Robert C. Davis, *Testing the Effectiveness of Batterer Programs and Judicial Monitoring*, Center for Court Innovation (November 2005).

# Chapter 1: California's Batterer Intervention Systems

## ***Introduction***

Every year in California over 100,000 arrests are made for misdemeanor and felony domestic violence charges.<sup>1</sup> Since 1994, California law has required defendants who are convicted and granted probation in these cases to complete a certified batterer intervention program (BIP).<sup>2</sup> In addition, recognizing the severity of the problem of intimate-partner violence and the unique challenges these cases present, many superior courts in California have adopted specialized procedures for handling domestic violence cases such as using dedicated calendars and holding periodic review hearings with probationers.

Adopting specialized procedures for handling domestic violence cases generally requires that courts coordinate their activities more closely with other justice system partners. Law enforcement, district attorneys, public defenders, the courts, probation departments, BIPs, victim-assistance programs, and other social service providers compose a *batterer-intervention system*. Working together, they form the system that confronts batterers with a variety of potential sanctions—ranging from incarceration to intensive monitoring by probation and the courts—as well as a requirement for rehabilitation through mandatory counseling and educational programs designed to change the attitudes and behavior of batterers.

Despite the clear interdependence of different justice system partners in the monitoring of domestic violence offenders, research on the efficacy of the justice system response to domestic violence has historically focused on individual components of the system. As a result, while arrest policy, domestic violence court monitoring, and BIP treatment modality have all been studied to determine the impact of these interventions, it remains unclear which elements of the system—sanctions, judicial review, frequency of review, intervention program modality, or some combination of these and other factors—ultimately reduce the likelihood of further violence by the batterer.

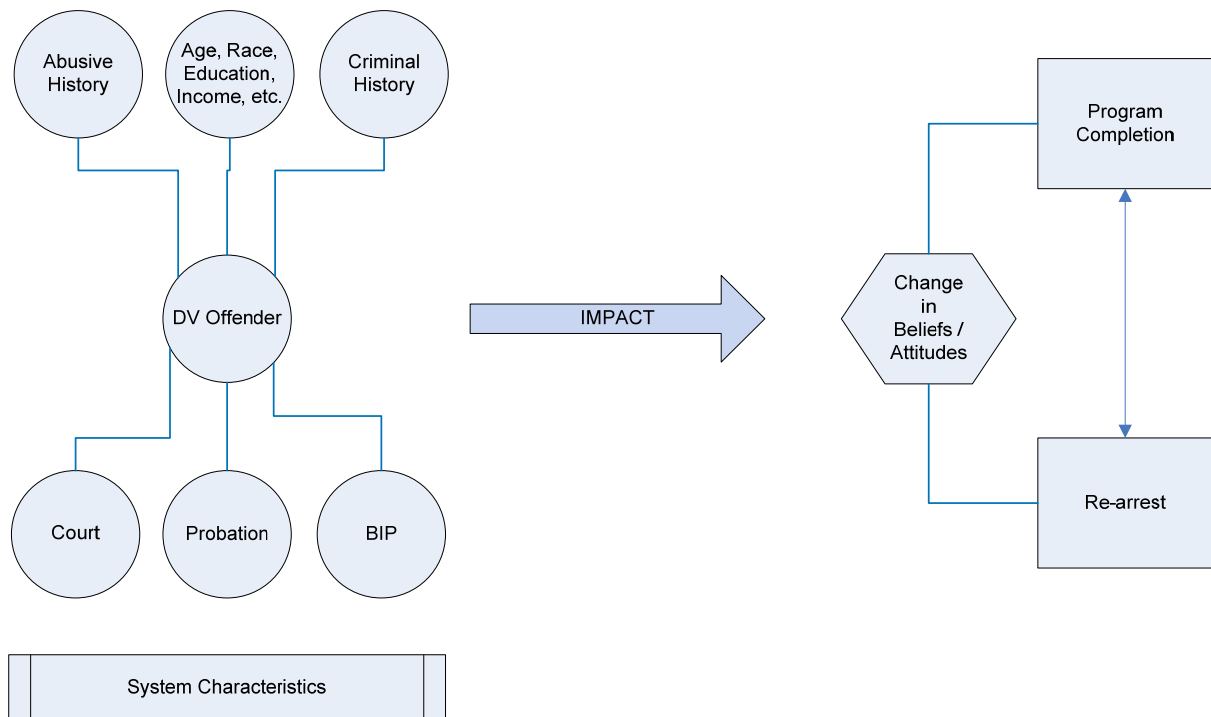
This study takes a systems perspective in evaluating the oversight of domestic violence offenders in five counties in California. The goal is not to study the effectiveness of these different jurisdictions per se, but rather to specify the system components and collaborative relationships among justice system partners that are most likely to improve compliance with court-ordered treatment programs and reduce re-offense in domestic violence cases. This study looks to document the differences that exist across jurisdictions and to understand the combined effect on domestic violence offenders of court, probation, and BIP oversight in different jurisdictions.

California's large population makes the state's justice system particularly well-suited for this type of evaluation. The large number of participants attending batterer intervention programs in the state made it possible to study a sample that provides greater confidence in the significance of the findings than in previous studies. Recent studies by Davis et al. and Feder and Forde<sup>3</sup> used total sample sizes of 376 and 404 respectively. In a four-site, cross-state evaluation conducted by Gondolf, the total sample size is 840.<sup>4</sup> This report draws on a sample of over a thousand men enrolled in treatment programs in five jurisdictions in the state.

In addition, the diversity of both case processing and BIP treatment models within a common legal framework makes it possible to hold a number of important factors constant while studying the effects of a range of other variables in the operation of the batterer intervention system. Regularly scheduled review by the court and the frequency of that review, frequency of oversight by probation, BIP treatment modality, and coordination among justice system partners can all be evaluated for their effect on offender compliance with the terms of probation and likelihood of reoffense. Rather than use random assignment—which has proved problematic in previous studies of batterer intervention programs—this study uses a quasi-experimental design, taking advantage of existing variation in the monitoring of batterers in California.<sup>5</sup>

The methodology and operationalization of measures is discussed in more detail below, but in brief the study isolates specific components of the batterer intervention system to assess how differences in the system interventions affect outcomes for men who are in the system.<sup>6</sup> Figure 1-A presents the logic model of the study and refers to these components of the batterer intervention system as “system characteristics”; it shows the logical relationship between these system characteristics, the characteristics of batterers, and the outcomes that the justice system seeks.

**Figure 1-A. Logic Model of Study Including Principal Variables**



## ***Literature Review***

In their review of the literature on batterer intervention, Davis et al. organize this research into three categories. Early research is characterized by studies that were largely qualitative and

descriptive seeking a better understanding of batterers, victims, and how the justice system and intervention programs work. Typically these studies lacked comparison groups against which to evaluate the outcomes.<sup>7</sup> More recent work in the field focuses on comparative outcome evaluation and can be classified into those that are based on experimental design and those that are based on quasi-experimental design.<sup>8</sup>

Regardless of design and methodology, research into the effectiveness of court-ordered treatment for batterers and periodic review of batterers has produced mixed results on whether or not these programs reduce the likelihood of further violence from an abusive partner.<sup>9</sup> One of the most exhaustive studies to date evaluated four sites in four different states and concluded that “the success of the programs appears to be related to the intervention system as a whole, and the programs may be inextricably embedded in the larger system.”<sup>10</sup>

Perhaps the most important reason for the uncertainty regarding program effectiveness is the existence of vast differences in the design and implementation of batterer intervention systems across the country. Comparative studies that can hold very little constant across multi-site evaluations face an enormous challenge in disentangling the impact that different components of the system contribute to program effectiveness. Furthermore, differences in system components are usually operationalized and measured with insufficient detail. Comparison groups are often dichotomized on the basis of the presence or absence of certain components, making the program a black box.

A growing body of literature on implementation analysis points to the importance of looking into the black box. By identifying and measuring system components more carefully, it is possible to link process evaluation with outcome analysis.<sup>11</sup> This study takes advantage of existing variation in system components at the level of the courts, probation, and batterer intervention programs. Because these components operate within a common statutory framework, a number of important system-level variables are held constant.

## ***Research Design and Methods***

As noted above, this study is designed as a quasi-experimental evaluation of batterer intervention *systems*. Although the statutory framework governing the handling of domestic violence cases in California applies statewide, important differences across counties make it possible to identify and measure different case-processing practices within county jurisdictions—the system—and to evaluate the impact of these practices on outcomes for domestic violence offenders.

This section describes the operationalization of the research design, the methods used to evaluate different components of the system, data collection instruments, and the data set on which we conduct analysis in subsequent chapters. We use both qualitative and quantitative methods to describe the various components of the justice systems in the study while the outcome data is entirely quantitative.

## **Changes to Original Study Methodology**

### **Court and Probation Jurisdictions**

The initial study design proposed to construct a sample in which half of the jurisdictions used dedicated domestic violence calendars and regularly scheduled review hearings, and the other half did not. Other considerations included finding courts that represented both Northern and Southern California and identifying jurisdictions large enough to provide substantial numbers of domestic violence offenders for the study over the course of three months of intake. Based on these criteria, the Superior Courts of Contra Costa, Los Angeles, Riverside, San Joaquin, Santa Clara, and Solano Counties were invited and agreed to participate in the research study.

However, before data collection started, Contra Costa County decided not to participate. Although the superior court and probation department had both agreed to provide the necessary data, some directors of the BIPs in the county were reluctant to take part. Many BIPs were participating in another research study at the time, and several felt it would be too difficult to manage the additional workload required to participate in this study.

The effect of losing Contra Costa County as a study jurisdiction at first appeared to be negligible. Enrollment projections provided by BIPs that had agreed to participate in the other five jurisdictions suggested that there would be more than enough subjects to compensate for the loss of Contra Costa, so a decision was made not to recruit another study county.

### **Program Fidelity**

The original study proposal envisioned identifying the principal treatment model of BIPs in the study jurisdictions, developing quantitative measures of these treatment models, and assessing the fidelity of the programs' adherence to the models. Researchers and policymakers in health and education have long recognized the importance of measuring how faithfully intervention mechanisms are implemented at the program level in order to draw accurate conclusions about the impact of different models. Program fidelity has only recently been addressed in domestic violence research.<sup>12</sup>

In our evaluation of program models, however, it became clear that different intervention models as practiced in California are not distinct enough to allow for the clear categorization of programs, let alone for the measurement of program differences according to categories. Even those BIPs that self-identified as adhering to one model or another borrow heavily from various traditions in their curricula and teaching style in practice. This finding prompted the research team to adopt a new approach in assessing the substantive content of BIPs.

Rather than evaluate the fidelity of programs to models that could not be fully disentangled from one another, we administered a survey of program content and teaching approach to BIPs in our study jurisdictions. This survey was developed on the basis of program descriptions and syllabi provided by BIPs, extensive interviews with program directors, and a thorough review of literature on domestic violence intervention models. Using the review of this information, a Program Content Survey (PCS) was developed to assess the substantive content of BIPs by collecting information from the programs regarding the following:

- The importance that programs attach to different educational topics;
- The frequency with which various educational topics are addressed in group;
- The importance that programs attach to certain teaching strategies and techniques; and
- The frequency with which different teaching strategies and techniques are employed in group.

The administration, results, and analysis of the PCS are discussed in greater detail in Chapter 3.

## **System Characteristics**

Measures of system characteristics include descriptive, qualitative information on the court, probation, and BIPs as well as quantitative data related to the nature and type of system contact with domestic violence offenders. A qualitative overview of the operations of the court and probation in the five jurisdictions in this study are provided in Chapter 2. This information is based primarily on interviews with key informants from the courts and probation and with directors of BIPs in the study jurisdictions, but it also incorporates some elements of the quantitative data collected for the study.

Appendix A provides a partial list of interviews conducted for the qualitative overview. This list is partial because it cannot document all of the back and forth that was involved in collecting data and that often resulted in a conversation or comment that was subsequently incorporated into the descriptive overview. Phone conversations, in-person interviews, stakeholder review of preliminary findings, conversations at conferences with directors of BIPs and staff from departments of probation and the courts, and observation of group sessions all contributed to the “thick description” in Chapter 2.

Quantitative measures of justice system contact with domestic violence offenders are drawn from the case management systems of the courts, from the probation departments, and from the attendance records of the BIPs. In one case, the research team constructed a database to assist a county department of probation in collecting data. To collect the quantitative data, staff in the courts and departments of probation were given unique identifiers—usually the court case number, probation number, and/or Criminal Identification Information number—for individuals in the sample and asked to match their records against these identifiers.

Data drawn from the courts’ case management systems include two types of data for each of the offenders enrolled in the study: charging information and post-sentencing hearings scheduled and held. Data from the departments of probation include the number and type of contacts that offenders had with probation.

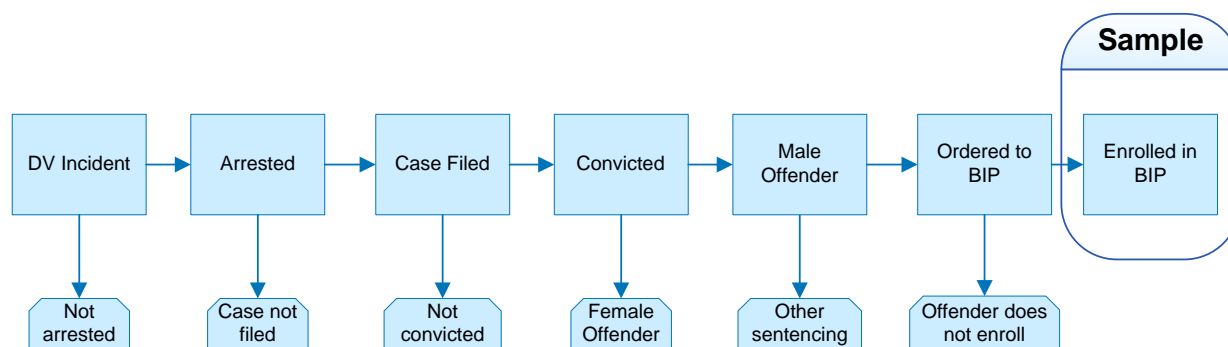
## Study Population

One key decision that was made to minimize variability across the sample was to limit the study to male offenders. Although domestic violence is perpetrated by and against both males and females, including offenders of both genders would have introduced yet one more major dimension along which the analysis would need to be stratified. Since the clear majority of domestic violence offenders who come before the courts are men, we focused the study on male offenders to capture the largest segment of domestic violence case processing by the justice system.

The study sample was further narrowed to include only men who were convicted of a *criminal* domestic violence offense against a female partner and ordered to attend a BIP as a condition of probation. Once again, this decision was made to limit the variability of the sample. Men may arrive at a BIP from a number of different paths including as a condition of a family court matter or even voluntarily. To increase the likelihood of capturing the effects of the justice system intervention, we chose to limit the variability of the underlying characteristics of the sample population.

Figure 1-B provides a graphic representation of the narrowing of the sample from the incidence of domestic violence in the general population through the justice system and into this study. We provide this diagram as a means of clarifying the scope of the research. As we show in Chapter 4, the characteristics of the study sample are not representative of either the population at large nor of domestic violence offenders generally. The fact that our sample does not reflect the broader social problem of domestic violence is a result of this winnowing process from domestic violence incident to enrollment in a BIP. As a result, conclusions that are drawn on the basis of this research need to be clearly delimited as applying only to those cases that fall into the far right-hand box in Figure 1-B.

**Figure 1-B. Narrowing of the Study Sample from Domestic Violence Incident to BIP Enrollment**



## Offender Characteristics: Participant Data Collected

Participating batterer intervention programs were asked to collect a variety of data on study enrollees. Many BIPs already collected very similar if not the same information on offenders as part of their intake process. To standardize these measures, however, we requested that BIPs collect the information on uniform data collection sheets, including:

- **Demographic Data**

The Supplemental Information Form (see Appendix B), administered at the intake session that each offender must complete before attending group sessions, was used to collect demographic data including education level; ethnicity; income; relationship to victim; family status; and whether the client had received counseling, had previously enrolled in a BIP, or had received treatment for drug/alcohol abuse or anger management. To help match the client data to court and probation records, the Supplemental Information Form also asked for the study enrollee's court case number and probation case number.

- **Revised Conflict Tactics Scale 2 and CAGE**

The Revised Conflict Tactics Scale 2 (CTS2) questionnaire (see Appendix C), filled out by the client at intake, seeks to assess behavior in the following areas: negotiation, psychological aggression, physical assault, injury, and sexual coercion. The questionnaire solicits responses to a series of statements about how the offender has dealt with disagreements with his partner over the last 12 months, with possible responses ranging from "never" to "more than 20 times." The questionnaire was modified slightly, with permission and exclusively for the purposes of this study, to include the CAGE assessment of alcohol abuse<sup>13</sup> and two questions regarding the respondent's current employment status. The instrument was translated into Chinese, Korean, and Spanish.

- **Criminal History**

In addition to the data collected by the programs, arrest history data was obtained from the California State Department of Justice (DOJ) for each enrollee. The DOJ database compiles information on arrests made by any law enforcement agency statewide. For this study, the database was queried for each offender's adult arrest history, including the date(s) of arrest, offense(s) charged, status of the offense(s), and disposition.

These data serve primarily as control variables, although the DOJ criminal history data also provides outcome data for tracking re-arrests. As control variables, this data allows us to isolate the impact of the intervention system on batterer behavior, providing greater confidence that the outcomes observed are not the result of spurious correlation with individual characteristics such as criminal history, age, or alcohol/drug dependence.

## **Outcome Data Collected**

The original study design proposed an outcome analysis based on two elements: program completion and re-arrest. To measure program completion, BIPs were asked to collect data on offender attendance in programs, including absences, termination, completion, and/or reenrollment, as applicable, and the date(s) of occurrence. For those who failed to complete the program, specific reasons for discharge (e.g., multiple absences, violation of probation, or re-arrest for any offenses) were recorded as well. For those programs without computerized attendance records, we developed an attendance data collection form to facilitate collection of this information (see Appendix D). Where we were able to provide assistance for data collection, members of the research team went on-site to BIP locations to assist in pulling case files and recording attendance data.



Attendance data may serve as both an outcome and a predictor variable depending on how it is used. For example, as an outcome variable, we might evaluate the individual characteristics that are correlated with longer periods of uninterrupted attendance; as a predictor variable, we might examine the effect on re-arrest of longer or shorter periods of uninterrupted attendance. Detailed attendance records, however, are less complete and their quality is less certain than data for the less-nuanced measures of program completion and termination.

Re-arrest data was collected from the same DOJ data download used to obtain an offender's criminal history. With cessation of further violence by the batterer as the ultimate goal of the batterer intervention system, re-arrest is defined simply as any record of arrest contained in the DOJ database including:

- Re-arrest for any offense (distinguishing between re-arrest for domestic violence and re-arrest for other crimes); and
- Violation of any probation conditions.

In addition to program completion and re-arrest data, one more outcome measure was added to the study: changes in the attitudes and beliefs of offenders. In the discussions regarding study methodology that took place as part of the BIP recruitment process, many BIP directors expressed concern that re-arrest and program completion data would not capture the more subtle effects of their programs. When it became clear that these programs may not have agreed to participate unless additional outcome measures were put into place, we then sought out a data collection instrument to capture more subtle changes in client behavior and attitudes as a result of participating in a BIP.

The instrument selected, called the BIP Process Survey (see Appendix E), is a questionnaire developed by Dr. Eric Mankowski of Portland State University in Oregon. It assesses psychosocial change and is composed of five subscales designed to assess a person's (1) sense of personal responsibility, (2) power and control beliefs, (3) understanding of the effects of abuse on others, (4) dependency on partner, and (5) anger control and management skills. Because this instrument seeks to measure change in the subject population, it needed to be administered twice during the time that clients were enrolled in the BIP—once approximately four weeks after intake and again just prior to program completion.

Though using the new questionnaire would require more work for BIPs, interestingly many directors—outside of those who had voiced concerns about the original outcome measures—expressed interest in administering this instrument to clients in their programs. As a result, the questionnaire was adopted on a voluntary basis for use in the study. BIPs in all jurisdictions sent data on this measure, although the size of the sample in Solano County is so small that the analysis of this measure excludes that jurisdiction.

## **Sampling Frame**

The study design called for tracking offenders from the point at which they enrolled in the BIP through a 6-month period *after* completing the 52-week program or from termination, as applicable. To ensure adequate time for post-program monitoring, a 3-month sampling frame

(April to July 2006) was established, and BIPs were asked to collect and submit data on all eligible clients who enrolled during the sample period.

## **Data Collected**

### **Constructing the sample**

Data collection commenced with a major recruitment effort to encourage BIPs to participate in the study. We sent information packets to certified BIPs in the five counties and held meetings to communicate directly with programs about the study and to encourage participation. Of the 155 certified BIPs in the five study jurisdictions, 73 agreed to participate in the study.

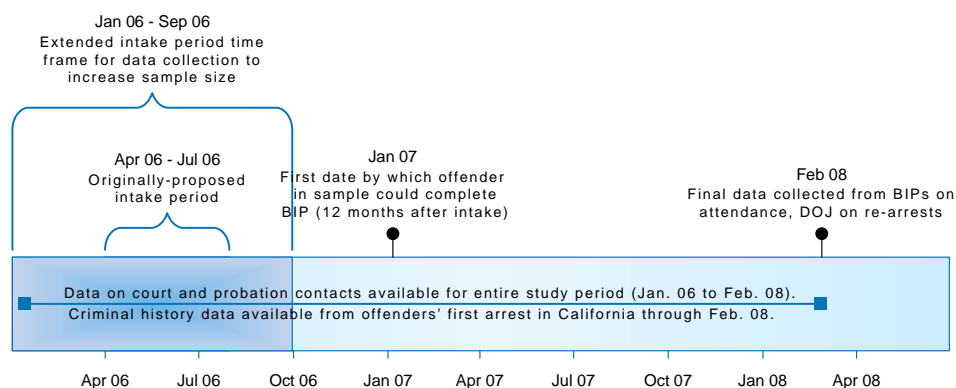
In addition to assessing the willingness of BIPs to participate in the project, we collected estimates of program enrollment to begin to estimate sample size. Based on these estimates, we expected the study to include more than 2,000 clients over three months of intake. As noted previously, even with the loss of Contra Costa County as a study jurisdiction and with only about half of the certified programs agreeing to participate in the study, the estimates provided by BIPs that had agreed to participate in the study actually exceeded the original DOJ estimates for the sample size of the study.

However, once data collection started, it became clear that the projections were overestimated. One cause might have been a misunderstanding of the client profile that was eligible to participate (i.e., parolees, referrals from Family Court, and offenders whose victims were not a current or former wife or girlfriend were among those excluded from the study). Another possibility is that the paperwork required to enroll and track the study participants became too burdensome to undertake for every eligible client, given that most BIPs operate with little to no administrative support staff. Newspaper articles in Los Angeles suggested that changes in law enforcement charging practices may have contributed to a decline in the number of domestic violence cases in that jurisdiction, but this does not explain the lower numbers that we saw across all of the study jurisdictions.<sup>14</sup>

Regardless of the cause, the lower-than-expected enrollment rates required us to shift tactics to increase the sample size. Project staff contacted and visited BIPs to encourage submission of client data and in some cases actually assisted with data collection from client records. Additionally, the intake period was extended two times from the initially designated period of April to July 2006. First, the sampling frame was extended to September 2006, with most BIPs agreeing to continue data collection for the additional two months. When that extension failed to produce enough enrollments to meet the projected study population of 2,000, the intake period was extended again for a group of BIPs that had already submitted some client data. Those BIPs were asked to submit basic descriptive data for clients enrolled as early as January 2006.

Figure 1-C shows the original time frame established for data collection; the extended time frame that was used to capture a large enough sample for the study; and a number of different relevant time frames for the analysis, including the minimum and maximum amount of time for which follow-up data was available on offenders based on enrollment.

**Figure 1-C. Study Time Frame**



By extending the time frames for program enrollment and through our enhanced recruitment efforts, we were able to compile an initial sample size of 1,457 clients enrolled in 53 BIPs. Table 1-A, below, shows the breakdown of the total sample by jurisdiction and also shows how complete the data is on different data measures. For example, although we received case numbers for 1,457 clients, supplemental information—demographic data, employment status, and relationship to the victim—was submitted for only 1,425 clients. While the fundamental information on clients needed to properly control for individual characteristics is relatively complete, the BIP Process Survey used to measure attitudes and beliefs was completed by only 685 clients following intake, and both pre- and post-program results are available for only 233 men in the sample.

All three data collection instruments for offenders—the CTS2, the Supplemental Information Form, and the BIP Process Survey—were administered on paper by the BIPs and returned by mail to the research team. Data was then entered from these forms into an Access database created for study.

**Table 1-A. Sample Size on Variables for All Data Collected**

|                                                  | Los Angeles | Riverside | San Joaquin | Santa Clara | Solano | Total |
|--------------------------------------------------|-------------|-----------|-------------|-------------|--------|-------|
| Offenders Enrolled in BIPs                       | 511         | 183       | 272         | 403         | 88     | 1,457 |
| CTS2                                             | 432         | 169       | 223         | 340         | 73     | 1,237 |
| Supplemental Information                         | 499         | 179       | 269         | 390         | 88     | 1,425 |
| BIP Attendance Records                           | 471         | 161       | 272         | 396         | 86     | 1,386 |
| BIP Program Completion/Termination <sup>15</sup> | 490         | 165       | 272         | 403         | 88     | 1,418 |
| All of the Above                                 | 410         | 151       | 223         | 340         | 73     | 1,197 |
| Number of Certified BIPs                         | 115         | 18        | 6           | 10          | 6      | 155   |
| BIPs Sending Client Data                         | 27          | 8         | 5           | 9           | 4      | 53    |
| BIP Process Survey—Intake                        | 232         | 91        | 143         | 207         | 12     | 685   |
| BIP Process Survey—Completion                    | 93          | 37        | 53          | 109         | 10     | 302   |
| Both Pre- and Post-BIP Process Surveys           | 74          | 30        | 32          | 90          | 7      | 233   |
| Matched with BIP Records                         |             |           |             |             |        |       |
| Court Docket Records                             | 459         | 168       | 266         | 390         | 81     | 1,364 |
| Probation Supervision Records <sup>16</sup>      | 19          | 57        | 271         | 384         | 28     | 759   |
| DOJ Arrest History Records                       | 434         | 156       | 254         | 387         | 72     | 1,303 |

A significant amount of time was devoted to tracking and logging attendance data. Obtaining this information was relatively easy from those programs with computerized attendance records but much more time-consuming for programs that track attendance manually. In some cases, project staff traveled to BIPs to record attendance information on study enrollees because programs could not spare the time or staff to do so.

Obtaining reliable termination and completion data for the study sample proved more challenging than anticipated. While we obtained information on the final status of 1,418 of the 1,457 enrollees in the study (see Table 1-B below), it is only for the 687 individuals who completed the study that we can consider those outcomes to be final. Those whose final status was listed as “terminated,” “active,” or “terminated and reenrolled” may have had a subsequent change in status. For example, an active or reenrolled individual could later be terminated or complete the program. A person whose status was recorded as “terminated” may have actually reenrolled in another BIP, with his status improperly recorded as “terminated” when it should have been “terminated and reenrolled.”

**Table 1-B. Final Status of Enrollees as of February 2008**

|                                            | Number       | Percent     |
|--------------------------------------------|--------------|-------------|
| Completed                                  | 687          | 47%         |
| Terminated                                 | 569          | 39%         |
| Active                                     | 70           | 5%          |
| Terminated and Re-enrolled                 | 92           | 6%          |
| No Data (BIP could not locate client data) | 39           | 3%          |
| <b>Total</b>                               | <b>1,457</b> | <b>100%</b> |

In some jurisdictions the courts might have been able to fill in the missing information regarding termination and re-enrollment; however, even if recorded, the data is not kept in an easily accessible format. While scheduling and occurrence of post-sentencing hearings are recorded in a court's case management system, details such as the name of the BIP are usually recorded in a text field in the register of actions. Tracking details such as whether or not an offender who was terminated re-enrolled in a program that was not participating in the study was not possible given the large sample size of this study. The availability and format of the termination and completion data impacted the analyses that could be undertaken in this study; further details are included in Chapters 4 and 5.

## **Court and Probation Data**

Project staff coordinated closely with Information Systems staff in the courts and probation departments to establish protocols for the matching, collection, and transfer of study data. A similar process was undertaken to secure arrest history data from the DOJ. Multiple rounds of testing were undertaken to make sure that the data could be matched from one source to another. Courts were requested to provide, for each offender enrolled in the study, the charges levied and the hearings held, specifically for the case that resulted in the referral to the BIP.

Charge data is based on a uniform DOJ code hierarchy, making this data relatively easy to work with across jurisdictions. In contrast, data on court hearings and probation contacts with offenders presented more of a challenge. There is no single statewide case management system for the California courts or for county departments of probation. As a result, each court and department of probation maintains its own unique database and corresponding coding system. This means that what is coded in one court as a "Proof of Enrollment" hearing may be called a "Probation Hearing Re: Enroll 52 Week Batterers Pgm" in another court. To further complicate matters, the text field to enter information on a hearing is usually a freeform field, meaning that the same "Probation Hearing Re: Enroll 52 Week Batterers Pgm" may also be referred to as a "Probation Hearing Re: Enroll 52Wks" or "Probation Hearing RE: Enroll 52WK BATTERER." Because each of these entries is worded slightly differently, they initially appear as three different types of hearings even though they seem to have the same purpose.

For purposes of analysis, the hearings data was consolidated into a common set of codes. Project staff collaborated with court staff familiar with criminal case data entry to determine how best to consolidate the numerous different codes into the following 10 hearing types, focusing just on post-sentencing hearings (See Table 1-C). In the analysis, this data was further collapsed to examine the frequency of contact with the court, distinguishing primarily between those contacts that are the result of an offender's violation of the terms of probation—such as arraignment on probation violation/warrant and bench warrant hearing—and those hearings that are held as part of a court's ongoing monitoring of offenders—such as progress report, proof of enrollment in program, and review hearing.

**Table 1-C. Consolidated Codes Used for Hearings Data Analysis**

| <b>Code</b> | <b>Description</b>                         |
|-------------|--------------------------------------------|
| AVP         | Arraignment on Probation Violation/Warrant |
| BWH         | Bench Warrant Hearing                      |
| PCK         | Probation Check/Hearing                    |
| PGR         | Progress Report                            |
| POC         | Proof of Completion                        |
| POE         | Proof of Enrollment in Program             |
| PVH         | Probation Violation Hearing                |
| RIN         | Reinstate Domestic Violence Program        |
| RWH         | Review Hearing                             |
| SVP         | Sentencing on Probation Violation          |

## ***Summary***

In this chapter we provide an overview of the research design and methods used for this study. We describe the system and offender characteristics in terms of the data that was collected from courts, probation departments, and batterer intervention programs. We also describe how the offender data sample and sampling frame were constructed, and explain some of the obstacles encountered in assembling this data. In the next chapter we delve into the system characteristics more fully with a qualitative description of the batterer intervention systems in the five study counties.

## Endnotes Chapter 1

---

1. California State Department of Justice, Criminal Justice Statistics Center, Monthly Arrest and Citation Register (MACR) (2003).
2. Pen. Code §1203.097 specifies that for a person “granted probation for a crime in which the victim is a person defined in Section 6211 of the Family Code, the terms of probation shall include” among other requirements “successful completion of a batterer’s program.” Fam. Code §6211 defines “Domestic Violence” as “abuse perpetrated against any of the following persons: (a) a spouse or former spouse. (b) a cohabitant or former cohabitant, as defined in Section 6209. (c) a person with whom the respondent is having or has had a dating or engagement relationship. (d) a person with whom the respondent has had a child, where the presumption applies that the male parent is the father of the child of the female parent under the Uniform Parentage Act (Part 3 commencing with Section 7600) of Division 12). (e) a child of a party or a child who is the subject of an action under the Uniform Parentage Act, where the presumption applies that the male parent is the father of the child to be protected. (f) any other person related by consanguinity or affinity within the second degree.”
3. R. C. Davis, B. G. Taylor, and C. C. Maxwell, *Does Batterer Treatment Reduce Violence? A Randomized Experiment in Brooklyn*, final report to the National Institute of Justice, Washington, DC (2000); L. Feder and D. R. Forde, *A Test of the Efficacy of Court-Mandated Counseling for Domestic Violence Offenders: The Broward Experiment*, final report to the National Institute of Justice, Washington, DC (2000).
4. E. W. Gondolf, *Batterer Intervention Systems: Issues, Outcomes, and Recommendations* (Thousand Oaks, CA: Sage Publications, 2002).
5. For a discussion of the methodological limitations of experimental evaluations see E. W. Gondolf, “Limitations of Experimental Evaluation of Batterer Programs” (2001) 2(1) *Trauma, Violence and Abuse*.
6. Although a fraction of the domestic violence cases that come through the system involve female batterers, the number of these cases is too small and the variation that they would introduce into an evaluation such as this too great to include them as part of the sample. The explicit focus of this study is men who are convicted of battering women.
7. E. W. Gondolf, *id.* note 3. Examples of research in this category include A. DeMaris and J. K. Jackson, “Batterers Reports of Recidivism After Counseling” (1987) 68 *Social Casework* 458–465; J. L. Edleson and R. J. Grusznski, “Treating Men Who Batter: Four Years of Outcome Data from Domestic Abuse Project” (1988) 12 *Journal of Social Service Research* 3–22.
8. Major studies based on experimental design include F.W. Dunford, “The San Diego Navy Experiment: An Assessment of Interventions for Men Who Assault Their Wives” (2000) 68 *Journal of Consulting and Clinical Psychology* 468–476; two studies funded by NIJ, one in New York City and the other in Broward County, FL, *id.* note 3; and a current NIJ-funded study in New York, M. Rempel, “Testing the Impacts of Court Monitoring and Batterers’ Intervention Programs,” unpublished progress report (2005). Numerous quasi-experimental studies have been conducted including E. W. Gondolf, *id.* note 4; D. G. Dutton, “The Outcome of Court-Mandated Treatment for Wife Assault: A Quasi-experimental Evaluation” (1987) 1 *Violence and Victim* 163–175; R. Dobash. et al., “Reeducation Programs for Violent Men: An Evaluation” (1996) 46 *Research Findings* 1–4.
9. Studies showing success in batterer intervention programs include R. M. Tolman and L. W. Bennet, “A Review of Qualitative Research on Men Who Batter” (1990) 5 *Journal of Interpersonal Violence* 87–118; S. H. Lund, N. E. Larsen, and S. K. Schultz, “Exploratory Evaluation of the Domestic Abuse Project,” in L. Ohlin and M. Tonry, eds., *Family Violence* (University of Chicago Press, 1989). A positive evaluation of domestic violence courts can be found in A. R. Gover, J. M. MacDonald, and G. P. Alpert, “Combating Domestic Violence: Findings from an Evaluation of a Local Domestic Violence Court” (2003) *Criminology and Public Policy*. Studies that raise questions about the effectiveness of batterer intervention programs include R. C. Davis, B. G. Taylor, and C. C. Maxwell; Feder and Forde, *id.* note 3.
10. *Id.* note 8.
11. A. Browne and A. Wildavsky, “Should Evaluation Become Implementation,” in A. J. Love, ed., *Developing Effective Internal Evaluation: New Directions for Program Evaluation* 1983 (20) 101–103; A. J. Love, “Beyond the Black Box: Practical Methods for Evaluating Implementation,” in J. S. Wholey, H. P. Halty, and K. E. Newcomer, eds., *The Handbook of Practical Program Evaluation*, 2nd ed. (San Francisco: Jossey-Bass, 2004).
12. E. Bowen and E. Gilchrist, “Comprehensive Evaluation: A Holistic Approach to Evaluating Domestic Violence Offender Programmes,” *International Journal of Offender Therapy and Comparative Criminology* (2004) 48(2).

- 
13. CAGE is an acronym formed by taking the first letter of key words for each of the four questions of the assessment (felt like you should cut down on drinking or drug use; felt annoyed by others criticizing your drinking or drug use; felt guilty about drinking or drug use; ever had an eye opener to steady nerves or get rid of a hangover).
14. *L.A. Daily News*, “‘Fudged’ Crime Stats Hide Domestic Abuse” (May 9, 2006).
15. We were not able to obtain final outcome data for 39 clients.
16. The low number of probation supervision records collected in some jurisdictions reflects the practice in these counties of assigning low-level criminal offenders to court-supervised or “informal” probation rather than formally supervising these probationers.





## **Chapter 2: Five Batterer Intervention Systems in California**

### ***Introduction***

California state law appears to create a uniform statewide system for the processing of misdemeanor and felony domestic violence convictions. Penal Code §1203.097 defines the terms of probation with which men convicted of domestic violence offenses are required to comply.<sup>1</sup> This section of the penal code includes specific provisions related to length of batterer intervention programs, size of groups, contents of BIP curricula, training of program staff, coordination with other justice system partners, and requirements for certification by county departments of probation. Indeed, the details contained in statute present such a strong appearance of standardization that some judicial officers have voiced their concerns about the application of a “one-size fits all” approach to the processing of domestic violence cases.

The apparent uniformity created by Pen. Code §1203.097, however, belies the operational reality of domestic violence case processing. Departments of probation, prosecuting attorneys, and public defenders are all part of local government—mostly county but sometimes city—and often operate quite differently from one jurisdiction to another. Until 1998, California’s courts were also administratively integrated into county government, and the legacy of unique local practices remains. As a result, the justice system response to domestic violence can vary considerably across and sometimes even within a single superior court jurisdiction.

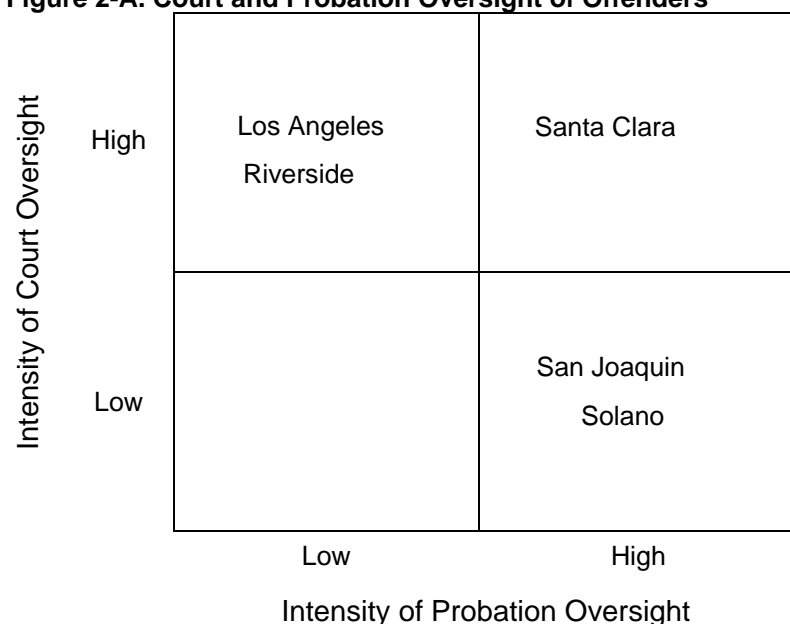
Differences in the ways that probation, courts, and law enforcement handle domestic violence cases from one county to the next may be further accentuated by differences in the BIPs themselves. Once again, although the penal code outlines relatively standard program content and format across the state, local variation is the rule, not the exception. Within the parameters established by Pen. Code §1203.097, there is considerable latitude for variability in BIP operations in terms of intervention strategies, background and training of facilitators, and operational capabilities for working with domestic violence offenders.

Understanding the effects of batterer intervention systems on men who are convicted of domestic violence crimes, therefore, requires that we understand differences across and within study jurisdictions that might influence the outcomes for men attending BIPs. In the following pages, we provide a qualitative overview of case-processing practices in the five counties from which our sample is drawn.

### ***Domestic Violence Case Processing in Five California Counties***

As noted in Chapter 1, we selected jurisdictions for this study with an eye toward capturing variation in court and probation oversight of domestic violence offenders and to provide a sufficiently large sample size to conduct statistical analysis of these variations. Figure 2-A, below, provides an overview of how the five systems in this study rank along two dimensions identified at the outset as critical to the justice system response to domestic violence: intensity of court oversight and intensity of probation oversight.

**Figure 2-A. Court and Probation Oversight of Offenders**



The different systems in the study are placed within these four quadrants on the basis of data drawn from court and probation records. The location of the different jurisdictions on the horizontal axis, representing intensity of probation oversight, is based on the frequency of offender contact with probation shown in the shaded cells of Table 2-A. While probation departments matched a little more than 50 percent of the total number of offenders in our sample, the range across different jurisdictions shows considerable variation, from a low of about 4 percent in Los Angeles to more than 99 percent of records matched in San Joaquin.

**Table 2-A: Probation Records and Offender Contact with Probation**

|                                                                                                                                         | Los Angeles | Riverside | San Joaquin | Santa Clara | Solano | Total |
|-----------------------------------------------------------------------------------------------------------------------------------------|-------------|-----------|-------------|-------------|--------|-------|
| Total Study Sample Size                                                                                                                 | 511         | 183       | 272         | 403         | 88     | 1,457 |
| Probation Supervision Records Matched (N)                                                                                               | 19          | 57        | 271         | 384         | 28     | 759   |
| Probation Supervision Records Matched (%)                                                                                               | 3.7%        | 31.1%     | 99.6%       | 95.3%       | 31.8%  | 52.1% |
| Average Number of Probation Contacts Per Month During First 3 Months After Intake (applied only to those offenders on formal probation) | 1.03        | 1.24      | 0.90        | 0.67        | 2.55   | 0.88  |
| Average Number of Probation Contacts Per Month During First 3 Months After Intake (applied to entire sample)                            | 0.04        | 0.39      | 0.89        | 0.64        | 0.84   | 0.46  |

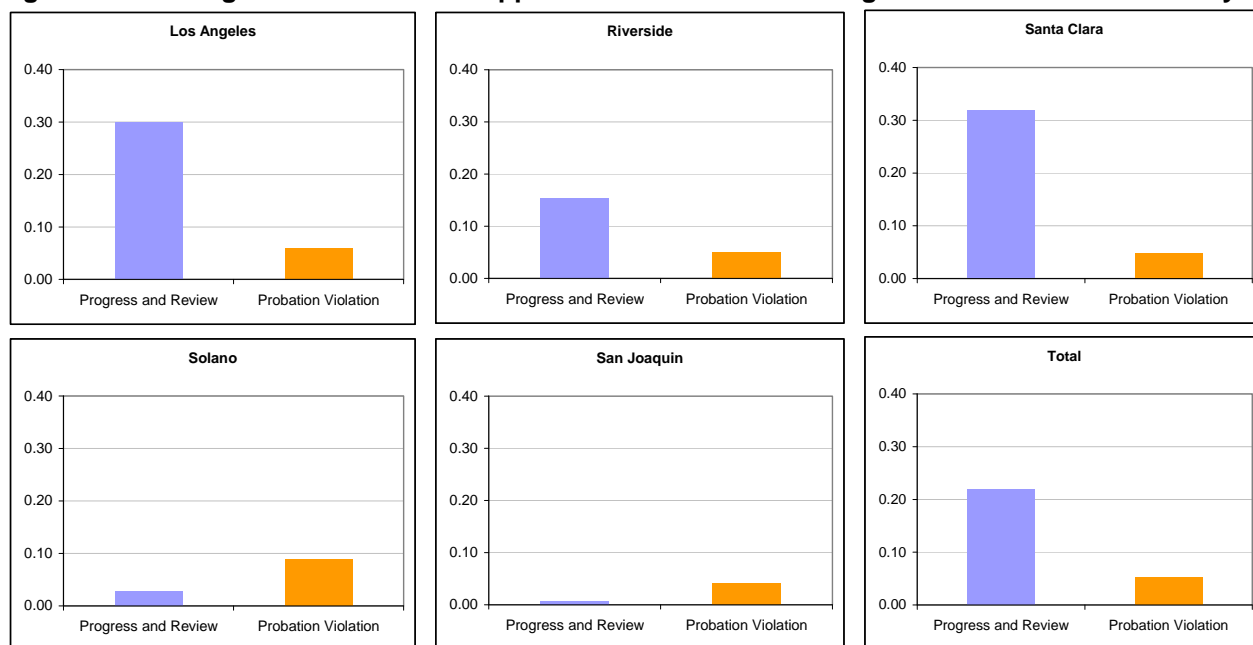
Matching of records provides one measure of oversight in that it corresponds roughly to the number of offenders who are on formal supervision by the department. Two other measures that can be used to estimate the intensity of probation oversight are (1) the average number of contacts with offenders who are on formal probation and (2) the average number of contacts for the entire sample. The latter measure—average number of probation contacts applied to the

entire sample—is used for locating jurisdictions on the horizontal axis because it appears to be a better representation of the intensity of probation oversight for the *system* as a whole. However imperfectly, this measure captures information about all of the offenders in the sample, not just those on formal probation. Looking at probation contacts only for those offenders who are on formal probation would inflate the estimate of probation supervision for a jurisdiction with a very low number of offenders on formal probation. For example, for the 19 offenders who appear in the probation database in Los Angeles, the level of oversight is actually quite high.

Jurisdictions are located along the vertical axis, representing intensity of court oversight, based on the data shown below in Figure 2-B. This data shows court appearances by offenders in the sample, distinguishing between court appearances that result from an offender violating the conditions of probation and those court appearances that are part of the court’s ongoing oversight of the offender, shown here as progress and review hearings.

The bars in Figure 2-B represent the average number of court appearances per month during the first three months following intake. Once again, the variation along this dimension is clear. The offenders in the sample from Santa Clara and Los Angeles have an average of almost one-third of a hearing per month for progress and review alone, suggesting that on average every offender in these jurisdictions returns to court once during the first three months following intake.

**Figure 2-B: Average Number of Court Appearances Per Month During First Three Months in Study**



## BIP Attendance Policy

Another type of offender oversight is the absence policy adopted in each jurisdiction. Although Pen. Code §1203.097(a)(6) states that any absence from the BIP without good cause or three absences with good cause require a court referral,<sup>2</sup> attendance policies in most if not all jurisdictions in the state appear to be more lenient than state law specifies. These policies came under criticism in a report to the Attorney General of California and in a Bureau of State Audits

(BSA) report discussed later in this chapter. According to the BSA report, the department of probation in San Joaquin County allowed as many as seven absences before terminating an offender from the BIP, while the departments of probation in Los Angeles, Riverside, and San Joaquin Counties all allowed makeup sessions for excused absences.

The BSA report had a clear and possibly measurable effect on the practices of the San Joaquin department of probation, which we discuss below. However, the focus of the BSA report on probation department policy may obscure underlying, intra-system variability because the frontline responsibility for monitoring and enforcing attendance policy resides with the BIPs. For example, while the BSA report points to a Los Angeles probation department policy of allowing three absences, this official policy is largely irrelevant to 96 percent of the men in our sample from that county because they are not formally supervised by probation.

The specifics of attendance policies reflect a commonality among jurisdictions—generally being more lenient than state law allows—and some differences that we do not fully understand. Important issues that we have not fully captured but that probably make a difference in the batterer intervention system’s impact include how decisions are made regarding termination, allowances for makeup classes, and the level of support that BIPs receive from the courts and probation in enforcing specific policies.

Table 2-B shows one measure of attendance policy differences across systems captured at the level of individual BIPs. Looking at the entire sample, on average men who completed the program or were still enrolled in the BIP at the end of the data collection period had 3.2 absences. Consistent with our interviews of key informants and with the BSA report, however, we can see variability across jurisdictions. San Joaquin’s more lenient absence policy is reflected in a higher-than-average number of absences than other jurisdictions, while Santa Clara’s stricter policy is reflected in a lower-than-average number of absences.

**Table 2-B. Absences for Men Who Completed or Were Still Enrolled in BIP**

|                                    | <b>Los Angeles</b> | <b>Riverside</b> | <b>San Joaquin</b> | <b>Santa Clara</b> | <b>Solano</b> | <b>Total</b> |
|------------------------------------|--------------------|------------------|--------------------|--------------------|---------------|--------------|
| Average Number of Absences         | 3.0                | 3.3              | 4.2                | 2.9                | 3.4           | 3.2          |
| Number Completed or Still Enrolled | 185                | 55               | 110                | 188                | 28            | 566          |

Besides the systemwide measures of oversight described above, each jurisdiction has unique features and policies for processing criminal domestic violence cases that may influence offender outcomes. The following sections consist of qualitative descriptions of each of the study jurisdictions to provide additional detail regarding case processing that cannot be captured through broader categorizations of court and probation oversight. While these descriptions cannot capture every element of the justice system that is relevant to domestic violence case processing, the intention is to highlight system policy in a number of key areas related to domestic violence. Specifically, these areas include court policy for monitoring offender compliance with the terms of probation; the department of probation’s policies for monitoring

offender compliance and for certification and approval of BIPs; and points of collaboration among the court, probation, and BIPs in the different jurisdictions.

## **Los Angeles County**

Looking again at Figure 2-A, the principal differentiation among study jurisdictions that we capture quantitatively is the intensity of court and probation oversight. Los Angeles County shows a high level of court oversight and relatively little probation oversight, reflecting the fact that relatively few offenders are formally supervised.

The department of probation appears to function in a reasonably uniform manner across the entire county of Los Angeles; however, prosecuting attorneys may differ in their approach to domestic violence crimes from location to location in the county depending on the city boundaries within which cases are prosecuted. The court, while unified in its organizational structure as a single superior court since 2000, also still retains certain features of its former organizational structure under which it once operated. That earlier structure included 26 municipal court locations, 14 superior court locations, and 1 justice court.

### **Court**

The 580 offenders from Los Angeles County in our sample were processed in more than 30 different court locations in the county. The largest single group of men, 156, were processed in the Long Beach courthouse. Another 57 were processed in the downtown court location, and more than 20 offenders were processed in each of another six locations, accounting for 168 men in the Los Angeles sample. Five or fewer men each were processed in 11 different locations in Los Angeles.

Because our sample is weighted heavily toward Long Beach and because the Long Beach courthouse is well known for its domestic violence case processing, our coverage of case-processing practices is most reflective of this court. Judges from Downey and El Monte were also consulted, and the description of Los Angeles County draws on these interviews and on interviews with other justice system partners as well.

In Long Beach, domestic violence cases are assigned to a vertical calendar, meaning that they are handled by a single judge from arraignment through post-disposition. Following sentencing, domestic violence offenders in Long Beach are referred to the Public Health Office in the courthouse, where they are provided a list of BIPs from which to select. The Public Health Office then records the BIP that is selected and follows up with the BIP three weeks after sentencing to confirm enrollment in the program.

The court in Long Beach and at least two other locations in Los Angeles County schedule review hearings for one month after sentencing to confirm enrollment. Key informants reported that, by the time of the enrollment confirmation hearing, offenders often fail to enroll for a variety of reasons such as inability to pay or loss of a job. After admonishing the offenders to enroll, the court will generally set another proof-of-enrollment hearing. Once the offenders have enrolled,

the Los Angeles courts we contacted then set regular three-month review hearings. In Long Beach, BIP progress reports are hand delivered by defendants to the judge in sealed envelopes at the review hearing.

### **Probation**

The probation department in Los Angeles County reports that it tracks probationers' compliance only for *felony* domestic violence cases. The fact that the vast majority of the cases in this study are misdemeanor cases means that probation had effectively no contact with men in the study and did not monitor compliance with orders. Matching the records of men in the study with department of probation records confirms the extremely limited oversight role played by probation in Los Angeles, as shown in Table 2-A above. Of the 511 men enrolled in BIPs in the Los Angeles sample, the database from the department of probation supervision records matched with only 19 offenders.

Resources for the certification and monitoring of BIPs in Los Angeles are also scarce. The department of probation reports that it has two monitors to track approximately 130 programs in the county and to track compliance with continuing educational requirements for program facilitators. Our interviews with the department of probation indicated that this allows the department to visit each BIP about twice a year to review files and sit in on group sessions. It does not allow for monitoring every facilitator in the programs that employ multiple facilitators.

### **Collaboration Among justice system partners**

The level of coordination among justice system partners in Los Angeles is relatively low. There do not appear to be any formal, regularly scheduled meetings among justice system partners in Los Angeles County to coordinate domestic violence issues. Neither the West Covina nor Los Angeles Central courts—the two court locations other than Long Beach that contributed qualitative information about case-processing practices in the county—held regular meetings with justice system partners.

Although there do not appear to be regularly scheduled meetings of justice system partners in Long Beach either, at this location the district attorney and public defender are present at all hearings and enjoy a good working relationship with one another and the court. The department of probation, however, does not have the resources necessary to assign a probation officer to Long Beach or any of the other domestic violence courtrooms in Los Angeles County.

## Riverside County

Riverside County appears in the same quadrant of Figure 2-A as Los Angeles. While the court appears to provide a relatively high level of oversight of domestic violence offenders, the department of probation has relatively little contact with offenders.

### Court

Data from the Superior Court of Riverside County on the 183 men in the sample from this jurisdiction do not allow us to distinguish among the four different court locations where domestic violence cases are heard. The inability to distinguish among court locations may create challenges of interpretation because of differences in case-processing practices between downtown Riverside and the other locations. In downtown Riverside, domestic violence cases are assigned to a dedicated calendar and regularly scheduled review hearings are held for offenders at three-month intervals. The other locations in Riverside where domestic violence cases are heard do not hold regularly scheduled review hearings.

Because the average number of hearings for domestic violence offenders in Riverside, shown in Figure 2-B, combines the results from these different case-processing practices across locations, this average may not provide an accurate representation of the court's practice as a whole. Further analysis will need to distinguish between the offenders processed in the main location downtown and those in the outlying locations.

In terms of BIP referral and follow-up, Riverside has a unique system in place to ensure that men granted probation for domestic violence crimes show up at a BIP and continue to attend. A nonprofit entity, the Volunteer Center, operates in Riverside's mid-county and central locations and serves as an intermediary between the court, the offender, the BIP, and probation.<sup>3</sup> Men convicted of a domestic violence crime in Riverside and granted probation must report to the Volunteer Center for intake and assignment to a BIP.

Key informants in Riverside suggested that while the review calendar in the main location was valuable, they also believed that the value had been diminished somewhat recently by the establishment of relatively short-term assignments of judicial officers to that calendar. Due to heavy workload in the court the judicial officer assigned to the domestic violence calendar in the downtown location rotated approximately every six months during the data collection period. Riverside informants believed that this limits the effectiveness of the review calendar by creating discontinuity in the oversight of offenders.

Another nuance of the court review process in Riverside is that these hearings are not necessarily in-court appearances before the judge. In some cases the progress review will occur with the clerk receiving and approving the paperwork or with a probation officer reviewing the report in the hall. As one informant from Riverside commented, this mode of reviewing progress reports diminishes the "audience impact" of everyone in court seeing success.



## **Probation**

As noted above, in Riverside the Volunteer Center manages the intake and assessment of domestic violence offenders. In terms of offender monitoring, most probationers are misdemeanants sentenced to informal, court-supervised probation. The probation department assigns one probation officer to the domestic violence court to review progress reports for those who are informally supervised. During the study period, Riverside County had received a Violence Against Women Act grant that allowed them to fund two additional probation officer positions to supervise domestic violence offenders under formal probation. In the department's desert division, probation also assigns a clerical position to monitor offenders.

Riverside's department of probation is active in monitoring BIPs and serves a coordinating function among justice system partners. The county appears to be unique among the jurisdictions in this study by requiring BIP facilitators to be either registered interns or licensed therapists with the Board of Behavioral Sciences. According to an informant from the department of probation, within California only San Diego and Riverside Counties require this level of training for BIP facilitators.

Riverside also appears to be unique in publishing detailed standards for BIPs. A booklet of over 100 pages, the "Standards for the Intervention and Treatment of Court Ordered Domestic Violence Offenders" is published by the department of probation annually and provides information on applying to become a certified BIP; references to appropriate penal code sections; clarification of the domestic violence intervention standards including intake procedures, length of treatment, content of programs, guidelines related to communication with the courts and referring agencies, and more.<sup>4</sup> This publication reflects a clear vision within the department of probation about what the county is seeking in program curriculum and content delivery as well as an effort to ensure adherence to that vision.

## **Collaboration among justice system partners**

Riverside has a moderate level of coordination among justice system partners, with the department of probation generally orchestrating this collaboration. The probation department holds twice annual meetings of justice system partners. These meetings are broadly attended by BIPs, judges who sit on domestic violence assignments, and representatives from the offices of the district attorney and public defender, as well as representatives from parole and the Volunteer Center.

The meetings are for reviewing changes to the law, coordinating case management among justice system partners, and serving as a forum for the department of probation to reiterate its standards of practice to directors of BIPs. At a meeting attended by the members of the research team, justice system partners worked to resolve problems related to probationers failing to enroll in a BIP, clarified interpretations of the penal code, and discussed the problem of unpaid fees.

In addition to these biannual meetings, since 1997 the Riverside department of probation has organized an annual Inland Empire Domestic Violence Conference. The conference is attended by justice agencies and BIPs from Riverside and other jurisdictions. The meeting generally

involves presentations on recent research or changes in domestic violence law and also sponsors BIP training.

## **San Joaquin County**

San Joaquin County presents a sharp contrast from Los Angeles and Riverside Counties. In San Joaquin County, monitoring of offenders in criminal domestic violence cases is largely a function of the probation department while there is relatively little oversight of offenders from the court.

### **Court**

The superior court hears domestic violence in four locations—Stockton, Tracy, Manteca, and Lodi—and the court holds internal meetings on a monthly basis to coordinate the handling of domestic violence cases. These monthly meetings are held in addition to bimonthly meetings of the criminal bench.

Despite these meetings to coordinate the court's response to domestic violence, one informant from the court indicated to us that local court practices remain different across the different court locations in San Joaquin. For purposes of analysis these differences may be less important than differences in Riverside because about three-quarters of the sample from San Joaquin—263 of the 348 offenders—were processed at the downtown location, with the rest of the sample evenly distributed in Lodi, Manteca, and Tracy.

### **Probation**

San Joaquin County was one of the jurisdictions selected for case file review by the BSA for its 2006 audit of domestic violence case processing. As a result of that audit's findings, the San Joaquin department of probation revised a number of its case-processing practices. Most of those changes were implemented in and around February 2007, which was during the data collection phase of the present project. As a result, offenders in the sample were exposed to a blend of past and current practices.

In the absence of a dedicated review calendar for domestic violence offenses in San Joaquin, BIP referral and follow-up are the responsibility of probation. Following sentencing, probationers have two to three weeks to come to probation for intake and placement. The department of probation normally needs about this long to get the probationer's data into its system and set up the case file. Probation then conducts an intake with the offender including a risk assessment.

As noted above, the department of probation in San Joaquin County implemented a number of changes in the management of its caseload as a result of the BSA report. Prior to the BSA report, released in November 2006, probation officers had some latitude in choosing a course of action for offenders who were out of compliance with the terms of probation, especially in the area of program absences. Since the release of the report, probation has eliminated some of that discretion, transferring responsibility to the courts for making decisions regarding violations of the terms of an offender's probation.

The probation department regularly monitors BIPs, but department officials indicated that San Joaquin has had the same BIPs since 1996 so there is more emphasis on ongoing program monitoring rather than certification of new programs. The department recently revised and formalized its monitoring practices and guidelines for maintaining certification.

### **Collaboration among justice system partners**

In the downtown Stockton location, a representative from probation is in the court during the domestic violence calendar and provides information to the judge. Coordination between court and probation, however, varies by location within the county. Informants in San Joaquin County indicated that while coordination with the Stockton court worked well, coordination with outlying courts did not necessarily work as well.

The San Joaquin County probation department holds quarterly meetings with BIPs to discuss the “problem of the quarter.” Since November 2006, that has meant responding to the BSA report, but generally the issues involve standardization of policies. These meetings appear to be exclusively between the probation and BIPs without the involvement of other justice system partners.

## **Santa Clara County**

The justice system response to domestic violence in Santa Clara County appears to be among the most intensive and coordinated of the jurisdictions in our sample. The county ranks high on the intensity of oversight of domestic violence offenders by both the court and department of probation. Additionally, this jurisdiction appears to have the most restrictive policy concerning absences and credits issued for BIP classes previously taken: offenders who are terminated from a BIP receive no credit for previous classes, regardless of the number of classes completed at the time of termination. Moreover, the actions taken by the court and probation are closely coordinated with one another through a number of different institutions within the county.

### **Court**

Although the offenders in our sample were processed in four locations, about 82 percent of the sample—317 cases out of 388—were processed in the downtown San Jose location, where most criminal domestic violence cases in Santa Clara County are heard. The domestic violence court there was reorganized during the course of the study. Under the current arrangement, one judge oversees arraignments, settlements, and case assignment to two departments. One department hears preliminary hearings, court trials, and jury trials, while the other hears preliminary hearings, review hearings, sentencing hearings, and probation violation hearings.

Offenders are required to complete an orientation with the department of probation within 10 days of sentencing and are required to return to court within 30 days for a proof-of-enrollment hearing. (Previously the court had established proof-of-enrollment hearings at 90 days from referral because of difficulties getting copies of the police report to probation and BIPs.)

Following the proof-of-enrollment hearing in Santa Clara, review hearings are scheduled every 60 to 90 days until program completion. Although Pen. Code §1203.097 requires that BIPs submit progress reports on offenders every three months, in Santa Clara the court requires these every two months.

## **Probation**

In addition to the court's supervision, the department of probation also maintains a high level of oversight of domestic violence offenders. All offenders initially start out under formal supervision, though there are varying levels of supervision depending on the individual's score on a risk assessment instrument administered at probation intake. Probation maintains several types of specialized supervision caseloads, including Spanish-speaking clients, deaf/mute clients, and domestic violence offenders with co-occurring mental health disorders.

In terms of BIP oversight, according to the department of probation, fully certified programs are subject to preannounced visits annually, and conditionally certified programs (those applying towards certification) are visited once every six months.

## **Collaboration among justice system partners**

In addition to the high level of oversight by both the court and probation independent of one another, Santa Clara County also appears to have the most actively coordinated response to domestic violence among the jurisdictions in our study. The Domestic Violence Council (DVC) and two standing committees of the DVC—the Batterer Intervention Committee (BIC) and the Court Systems Committee (CSC)—all meet monthly to address different aspects of domestic violence in Santa Clara.

The DVC, established in 1991 as an advisory body to the board of supervisors, seeks to improve coordination among the court, members of the community, victims, and county agencies and departments. Monthly BIC meetings—attended by representatives of the court, probation, the district attorney's office, the public defender's office, and directors of BIPs—are held to exchange information needed to address offender accountability and victim safety. Monthly CSC meetings include the BIC attendees as well as representatives from the private bar, law enforcement, and other service providers (in addition to BIPs). The focus of the CSC is to improve the court's handling of domestic violence cases and to educate service providers so that they can assist victims in accessing and navigating the court system.

The Superior Court of Santa Clara County is actively engaged in coordinating its response to domestic violence both internally and in collaboration with other justice system partners. Internally, the Domestic Violence Coordinating Committee (DVCC) meets every other month to follow up on issues that arise in other forums. Domestic violence court judges also schedule quarterly meetings to meet with probation, BIPs, the district attorney's office, and the public defender's office.

Finally, a group organized by the court called Filling the Gaps provides an overview to the court and justice system partners to help coordinate domestic violence issues in the court. Filling the

Gaps meets every three to four months to discuss reports from various committees of the DVC—the research committee, firearms committee, and safety committee—and identify gaps and concerns related to domestic violence cases that cross different case types such as family, criminal, probate, and juvenile. To assist with the coordination of these cases the Superior Court of Santa Clara County recently hired a case manager to locate and track related criminal, family, and juvenile cases and ensure that judges are aware of restraining orders or other pending matters that may have been issued from a different division of the superior court.

## **Solano County**

Solano County most closely resembles San Joaquin County in terms of the level of court oversight (low) and probation oversight (high). Solano is also unique in that it offers a deferred entry of judgment plea to certain low-level, mostly first-time offenders. These cases are supervised informally by the courts rather than through the probation department.

### **Court**

Data provided by the court does not allow us to identify the particular location in which the offenders in the Solano sample were processed, but misdemeanor domestic violence cases are heard in two locations in the county: Fairfield and Vallejo. In both sites the cases are arraigned in a single department presided over by a commissioner. Felony cases are randomly assigned to a judge on the criminal bench.

Offenders are required to report to probation within two days of sentencing. Although there does not appear to be any follow-up hearing to confirm enrollment in the program—probation handles this—the court does track misdemeanor cases by setting review hearings at 6 months from program enrollment and at 18 months to verify program completion. Progress reports for men on informal probation are faxed to the court, while progress reports for men on formal probation are sent to the department of probation.

### **Probation**

The probation department conducts the intake and monitoring of offenders sentenced to formal probation. Two probation officers oversee specialized caseloads in the jurisdiction, including one who monitors cases where children witnessed the domestic violence and/or cases where there are family reunification issues. The other specialized caseload is for offenders who require intensive supervision. The offender is referred to a program at his intake with probation and then scheduled for a follow-up meeting in two weeks at which he is required to show proof of enrollment.

In Solano County the district attorney's office also plays a monitoring role similar to that of the department of probation. It has dedicated domestic violence prosecutors who follow cases and develop relationships with victims in order to monitor compliance. The prosecutors assigned to domestic violence cases can recalendar cases for a court hearing when they encounter problems with compliance and can charge offenders with violations as well as new charges.

With respect to certification and monitoring of BIPs in Solano County, the probation department employs a clinical services associate to conduct program monitoring, including observing facilitators and conducting file reviews.

### **Collaboration among justice system partners**

At our initial meeting with representatives of the department of probation in 2007, we learned that probation was in the process of implementing a new domestic violence database. The new system was designed to link BIPs directly to probation via a web-based program that would allow BIPs to transfer information on enrollment and progress directly into a database viewable by court and probation staff.

Probation meets quarterly with BIPs to reiterate the specifics of state law and county policy regarding these programs. It does not appear that the court or other criminal justice agencies are present at those meetings. The new domestic violence database, which will track offender enrollment and progress continuously, may be an important step in improving the sharing and exchange of information.

### **Other Issues**

In addition to these jurisdiction-specific features, two additional issues common to all five systems in the study are worth highlighting. They either made the practice of processing domestic violence offenders different from the letter of the law or may have changed local case processing during the course of the study.

### **Jail Overcrowding**

Informants in all the systems we studied cited jail overcrowding as a problem. The exact magnitude of the problem may not have been the same in all jurisdictions. And in one jurisdiction we learned that the implication of overcrowding—the reduced percentage of time offenders would actually serve if they opted for a jail sentence—changed over the course of the study. Nonetheless, interviews with representatives of courts, departments of probation, and BIPs in all of the study jurisdictions referred to jail overcrowding as a problem in providing credible sanctions to domestic violence offenders. Jail sentences were frequently reduced to ease jail overcrowding, and domestic violence offenders familiar with the justice system were said to opt out of the probationary term and choose jail time instead, knowing that the time served would be a fraction of the actual sentence.<sup>5</sup>

### **External Scrutiny**

A number of high-profile, statewide reports on the justice system response to domestic violence were released either during or around the time of this study. In some jurisdictions it was clear that a particular report had changed the practice of case processing, and we noted these changes in the descriptive overview of individual jurisdictions. Even where a direct impact of these

reports was not clear, the possible impact should be kept in mind as an external factor that may have altered the practices in any one of the jurisdictions during the course of this study.

- Data collection was initiated less than a year after the release of a report to the California Attorney General that was critical of law enforcement, court, and probation department responses to domestic violence.<sup>6</sup>
- Partly in response to the report to the Attorney General, the Chief Justice of the California Supreme Court appointed a task force in September 2005, to recommend changes to improve court practice and procedure in domestic violence cases. The task force worked throughout the study period and released its report to the Judicial Council of California in January 2008. As part of the work of the task force, the Judicial Council approved a new Batterer Intervention Program Progress Report form to be used by probation departments or BIPs to inform the courts of the progress of offenders enrolled in BIPs.<sup>7</sup>
- In November 2006, the Bureau of State Audits released a report that highlighted failures of county probation departments and the courts to comply with state law related to BIPs. The BSA report focused on practices in five California counties, including three counties that are part of the present study: Los Angeles, Riverside, and San Joaquin.<sup>8</sup>

## **Summary**

This chapter provides a qualitative overview of the batterer intervention system in each of the five study jurisdictions, describing relevant characteristics of the courts and probation departments and the level of collaboration among justice system partners. In addition to the county-specific characteristics, two issues common to all jurisdictions—jail overcrowding and external scrutiny—are also highlighted. Continuing this report’s thematic progression from overarching, system-wide traits to more individualized levels of analysis, the next chapter discusses characteristics of and findings related to the batterer intervention programs that participated in this study.

## Endnotes Chapter 2

---

1. California law applies equally to men and women. As noted in the previous chapter, our sample is restricted to male domestic violence offenders so we use the term “men” here both as a reflection of that fact and for the sake of simplicity.

2. *Keeping the Promise: Victim Safety and Batterer Accountability*, Report to the California Attorney General from the Task Force on Local Criminal Justice Response to Domestic Violence (June 2005), pp. 68–69; and *Batterer Intervention Programs: County Probation Departments Could Improve Their Compliance with State Law, but Progress in Batterer Accountability Also Depends on the Courts*, California State Auditor, Bureau of State Audits (November 2006), pp. 22–24.

3. The Riverside County Department of Probation has a clerical position in its desert location that monitors offenders. Although this discussion is focused on the role that the Volunteer Center plays in the management of the domestic violence caseload, the nonprofit agency provides referral and intake services for programs other than the BIPs. The Volunteer Center refers offenders to court-ordered programs for community service, anger management, parenting, and child abuse in addition to the BIP referrals.

4. *Standards for the Intervention and Treatment of Court Ordered Domestic Violence Offenders* (2008), Alan M. Crogan, Chief Probation Officer, Riverside County Probation Department.

5. A *Los Angeles Times* article cites Los Angeles County Sheriff Lee Baca’s estimate that “male inmates serve an average of 70% of their sentences.” While the exact percentage of time served by men convicted of domestic violence crimes is not known, in all of the study jurisdictions, the fact that few offenders serve the full length of their sentences was cited as problematic for creating effective sanctions for offenders. See “Los Angeles County Braces for an Influx of State Prisoners,” *L.A. Times*, (May 27, 2008).

6. *Keeping the Promise*, *id.* note 2.

7. *Recommended Guidelines and Practices for Improving the Administration of Justice in Domestic Violence Cases*, Final Report of the Domestic Violence Practice and Procedure Task Force (January 2008) and Judicial Council Form CR-168.

8. *Batterer Intervention Programs*, *id.* note 2.





## Chapter 3: Batterer Intervention Program Content

### ***Introduction***

Despite the measurable differences in court and probation oversight of domestic violence offenders, batterer intervention programs (BIPs) are probably the single most important component of the justice system's intervention in these cases in California. Even in jurisdictions with active monitoring by probation and regular review hearings by the court, the mandatory, weekly, two-hour BIP sessions give these programs, among all the justice system partners, the most continuous, direct contact with domestic violence offenders. While BIPs are an essential part of the justice system's response to domestic violence, they are in an equally important sense nested within the justice system. They are subject to state law regarding the form and content of their programs, and they must be certified annually by the county department of probation.

To better understand the impact of domestic violence case processing across the five jurisdictions in this study, we sought information about the content of the intervention programs. This chapter describes and evaluates data that the research team collected about the content of different BIPs in our study jurisdictions. Survey data on the importance that BIP directors place on various educational topics and skills training as well as on the frequency with which these are taught in group sessions indicates that there is little difference across jurisdictions in the reported content of programs. While this finding suggests that it may be possible to effectively hold the program content constant across jurisdictions, substantial variation in *outcomes* across BIPs within and across jurisdictions suggests something different. It may be that the content of the programs as described in these survey instruments is less important than the actual implementation of the content in group sessions.

In addition to finding very similar approaches across all of the study jurisdictions, the survey of program content found that BIPs across the entire sample have adopted integrative approaches to their intervention models. BIPs report employing educational models and skills training that include, at a minimum, elements of both the Duluth and cognitive-behavioral models (see Appendix F for an overview of each model.) These findings also indicate that programs tend to emphasize educational topics over skills training for batterers, suggesting that BIPs find it necessary to introduce program content in a way that is appropriate for the educational and developmental levels of their clients.

### ***Background***

The Program Content Survey (PCS) was developed to assess the substantive content of the BIPs participating in this study. The goal of this assessment was to provide the research team with information on the educational topics, coping skills, and teaching techniques that BIPs employ in their interventions with male offenders. On that basis, the PCS seeks to document the full array of elements that any intervention program might incorporate into its educational treatment program with the expectation that no single BIP would cover all of these materials or techniques.

The strategy for the development of the PCS was to first identify the principal sources of information pertaining to intervention models designed for male batterers. Two models for the treatment for abusive men figured prominently in this process: the Duluth model and the cognitive-behavioral model. Although numerous other sources of information were drawn upon to construct the PCS—including BIPs' own program descriptions and course syllabi, as well as articles and books focusing on intervention and treatment approaches to domestic violence—the influence of these two models is so pervasive that their components formed a significant proportion of the items contained in the PCS.

The development of the survey involved an iterative process, beginning with gathering information from the sources described above. The information gathered from these sources was at an intermediate level of generality, focusing on educational concepts and topics identified as important in a 52-week program, coping strategies and techniques training thought useful to help batterers end their abusive behavior, and the teaching strategies facilitators employed in their interventions with offenders.

After identifying these indicators, we narrowed the list to avoid unnecessary overlap and to achieve reasonable time limits for the administration of the survey. We developed additional survey items to assess whether a specific concept, coping strategy, or teaching technique was employed by a given BIP. As we produced drafts of this survey from the list of indicators and items, we circulated them among members of the research team. Clinical advisors working with participating BIPs reviewed the drafts once they became more advanced, and their feedback was integrated into subsequent versions of the PCS. We maintained this iterative process until we arrived at the present survey, which was then mailed to participating BIPs.

## **Administration Procedure for the PCS**

BIPs participating in the study received by mail the PCS along with an instruction sheet and relevant contact information for a member of the research team. After the initial mailing, we contacted BIPs through e-mail messages and/or phone calls to encourage completion of the survey and to answer questions related to the survey. The instructions for the PCS indicated that a senior group facilitator or program manager who was highly familiar with the intervention program curriculum and men's groups should fill out the survey. Respondents were encouraged to consult with other facilitators about the specifics of program elements as they thought necessary.

The instructions also encouraged respondents to think of one or two of their group facilitators (including themselves if appropriate) who best characterized how their program approached the use of its curriculum and intervention with batterers. They were then encouraged to use these facilitators as referents when responding to the survey. If ongoing consultation with these and other program facilitators was deemed helpful when filling out the PCS, respondents were encouraged to do so.

We provided decision criteria to respondents to help them select those educational topics, coping strategies, and teaching techniques that their programs covered. In addition, we provided

suggestions to respondents for determining how frequently program elements were covered during class sessions and for rating the importance of PCS items. Finally, we assured respondents that their personal identities as well as their organizations' identities would remain confidential. Of the 73 BIPs that received the PCS, 45 completed and returned the survey.

### **Caveats Concerning Responses to the PCS Data**

As the PCS instructions indicated, the subjects and topics covered by the survey represent a broad overview of what intervention programs with varying orientations might cover. We did not anticipate that a single intervention program could or should try to cover all of the program elements identified in the survey given the limited time and resources many programs have at their disposal. However, the majority of programs did indicate that they undertook teaching and training in most of the areas covered by the PCS.

This suggests that while programs appear quite ambitious about what they try to cover in their 52-week programs, there may be significant overestimation by many BIPs regarding the scope and intensity of the *formal* training they undertake with batterers. Where specifics of any tendency toward overestimation of program content and activity is not known, anecdotal information suggests that the more highly trained facilitators had a tendency to be more conservative in their estimates of what their programs undertake in terms of the formal curriculum.

The moderate response rate of 61 percent achieved in the administration of the PCS suggests that caution should be used in interpreting the findings to characterize intervention approaches for a given court jurisdiction. Further, jurisdictions vary greatly in the number of BIPs that participated in the present study, reflecting a number of factors including the tendency for counties with smaller populations to have proportionately fewer BIPs. The relatively small number of BIPs present in certain counties reduces the power of statistical tests and thus makes it difficult to detect reliable differences among jurisdictions in their approaches to batterer intervention.

As may also be seen in the forthcoming description of findings for the PCS, the very utility of court jurisdiction as a reliable way to group BIPs, in relation to their responses to the PCS, may be called into question. This suggests that there may be more useful ways to categorize BIPs in relation to their approaches to batterer intervention, and it constrains what can be concluded about jurisdictional differences.

## ***Findings***

### **Educational Topics**

The educational topics identified by survey respondents as important in helping their clients end their domestic violence appear quite consistent with state and local mandates calling for holding

batterers accountable for their domestic abuse. Further, those elements rated as being of higher importance by program staff appear central to the Duluth and cognitive-behavioral models described previously, although innovative approaches related to attachment and personality theories,<sup>1</sup> interpersonal communication, and community-cultural approaches are also reported. (See Table 3-A below for educational topics rated of highest importance and Table 3-A1 of Appendix G for all other subjects.)

### **Ratings of Importance: Educational Topics**

Across court jurisdictions, educational concepts commonly identified by the preceding models as important to successful intervention programs were frequently rated highly, including the importance of addressing:

- Accountability and personal responsibility
- Beliefs and attitudes that provide the basis for domestic abuse
- Stress management and effective coping
- Power and control in abusive situations
- Management of anger and emotion
- Understanding the effects of abuse

More specifically, survey items assessing BIPs' coverage of holding batterers accountable, the common defense mechanisms used by batterers to justify their abuse (including minimization and blaming, power and control issues in abusive relationships, anger and emotion management, attitudes and beliefs underlying abusive behavior, stress and coping, and alcohol and substance abuse) are all thought to be important to very important to cover during a 52-week program (Table 3-A). This indicates that BIPs may be taking a cross-disciplinary approach to the topics and issues they address in group sessions and/or through assignments that facilitators make during a 52-week program.

**Table 3-A. Importance Ratings for Educational Topics**

| Educational Topics Explained or Discussed |                                                                                                       |                                   | Average Rating of Importance, by Jurisdiction |                    |     |     |     |     |     |      |
|-------------------------------------------|-------------------------------------------------------------------------------------------------------|-----------------------------------|-----------------------------------------------|--------------------|-----|-----|-----|-----|-----|------|
| Item                                      | Topics and Issues                                                                                     | Coding Cat.                       | N of BIPs Covering Topic                      | Average Importance | LA  | RIV | SC  | SJ  | Sol | Sig. |
| Q2A                                       | Accountability and taking responsibility for domestic abuse                                           | Accountability                    | 43                                            | 4.8                | 5.0 | 5.0 | 5.0 | 5.0 | 4.8 | ns*  |
| Q17A                                      | Denial of abuse as defense mechanism                                                                  | Defense Mechanisms: Batterers     | 42                                            | 4.5                | 4.4 | 4.7 | 4.1 | 5.0 | 4.5 | ns   |
| Q21A                                      | Effects of abuse on children                                                                          | Abuse                             | 43                                            | 4.5                | 4.3 | 4.3 | 3.9 | 4.0 | 4.5 | ns   |
| Q35A                                      | Minimization of abuse as defense mechanism                                                            | Defense Mechanisms: Batterers     | 43                                            | 4.5                | 4.6 | 4.9 | 4.3 | 4.5 | 3.8 | ns*  |
| Q22A                                      | Effects of abuse on partner                                                                           | Abuse                             | 43                                            | 4.5                | 4.6 | 4.3 | 4.0 | 4.7 | 4.8 | ns   |
| Q43A                                      | Power and control dynamics in abusive relationships                                                   | Power & Control                   | 42                                            | 4.4                | 4.4 | 4.4 | 3.9 | 4.8 | 5.0 | ns   |
| Q50A                                      | Understanding the personal consequences of one's abusive behavior                                     | Abuse                             | 42                                            | 4.3                | 4.3 | 4.4 | 3.8 | 4.7 | 5.0 | ns   |
| Q29A                                      | Identification of abuse triggers (anger, fear, grief, loss, separation, jealousy)                     | Anger & Emotion Management        | 42                                            | 4.3                | 4.5 | 4.7 | 2.9 | 4.7 | 5.0 | ns*  |
| Q49A                                      | Time-out technique or procedure explained                                                             | Anger & Emotion Management        | 43                                            | 4.3                | 4.5 | 4.4 | 3.3 | 4.5 | 5.0 | ns*  |
| Q4A                                       | Anger and anger triggers                                                                              | Anger & Emotion Management        | 43                                            | 4.2                | 4.3 | 4.6 | 3.3 | 4.5 | 4.8 | ns   |
| Q6A                                       | Beliefs and attitudes leading to domestic abuse                                                       | Attitudes & Beliefs               | 43                                            | 4.2                | 4.5 | 4.0 | 4.0 | 4.0 | 4.0 | ns   |
| Q7A                                       | Blaming of others as defense mechanism                                                                | Defense Mechanisms: Batterers     | 41                                            | 4.2                | 4.4 | 4.7 | 3.9 | 3.8 | 3.5 | ns   |
| Q14A                                      | Conflict resolution techniques                                                                        | Conflict Resolution & Negotiation | 43                                            | 4.2                | 4.6 | 4.4 | 3.0 | 4.2 | 4.3 | ns*  |
| Q20A                                      | Domestic abuse: What is it behaviorally? (e.g., emotional, economic, sexual, isolation, intimidation) | Abuse                             | 43                                            | 4.2                | 4.3 | 4.3 | 3.9 | 4.0 | 4.5 | ns   |
| Q13A                                      | Cognitive restructuring                                                                               | Cognitive-Behavioral              | 42                                            | 4.2                | 4.5 | 4.3 | 3.5 | 3.5 | 4.8 | ns*  |
| Q53A                                      | Wheel of power and control in relation to domestic abuse                                              | Power & Control                   | 42                                            | 4.1                | 4.4 | 4.3 | 3.6 | 3.5 | 4.8 | ns*  |
| Q30A                                      | Identification of high-risk situations                                                                | Stress & Coping                   | 42                                            | 4.1                | 4.4 | 4.4 | 3.5 | 3.7 | 4.3 | ns*  |
| Q1A                                       | Accepting and working with victim's anger, resentment, and distrust as result of abuse                | Accountability                    | 41                                            | 4.1                | 4.4 | 3.7 | 4.3 | 3.7 | 3.5 | ns*  |
| Q41A                                      | Personal responsibility and honesty on an everyday basis                                              | Accountability                    | 42                                            | 4.0                | 4.2 | 4.0 | 3.9 | 3.3 | 3.5 | ns   |

Note 1: LA=Los Angeles, RIV=Riverside, SC=Santa Clara, SJ=San Joaquin, Sol=Solano

Note 2: ns=Nonsignificant differences among jurisdictions at the 5 percent level or beyond; ns\*=Nonsignificant differences with Tests for Homogeneity of Variance indicating that jurisdiction is not a reliable way to group BIPs for the corresponding item.

### **Jurisdictional Differences and the Importance of Educational Topics**

It is important to note that no significant differences were found among jurisdictions in terms of the average ratings with which they assigned importance to educational topics rated as important to very important for helping batterers. However, statistically significant differences were noted among a number of educational topics thought by respondents to be *less* important in ending domestic abuse (Table 3-A1 of Appendix G). They include the following educational topics, all significantly different at the level of 5 percent or better:

- Client's family of origin as a source of his attitudes and beliefs;
- Coping with separation and/or divorce from a partner;
- Cultural and societal norms supporting aggression against women and others;
- Healthy versus unhealthy relationships with a domestic partner;
- Racism as related to client's self-concept and attitudes to self and partner; and
- Effects of domestic abuse on other adults and the community.

Variation in the ratings of importance of these topics may be linked to a number of factors, including differences in views about the causes of domestic abuse among male batterers, the specific needs of local client populations, and differences among BIPs in their interpretation of local mandates for the treatment of abuse.

### **Jurisdiction as a Grouping Variable and the Importance of Educational Topics**

Grouping BIPs in accordance with the court jurisdiction in which they are located often does not appear to be a statistically reliable way of characterizing their ratings of the importance of educational topics in ending domestic abuse. This is largely because differences among BIPs within a given court jurisdiction are often larger than differences among jurisdictions.

For example, ratings of the importance of educational topics—including accountability for domestic abuse, minimization of abuse by batterers, anger and emotion management, and conflict resolution—all vary more among BIPs in the same court jurisdiction than across jurisdictions (Table 3-B, below, and Table 3-B1 of Appendix H). This suggests that differences among BIPs in terms of the emphasis they place on a number of important educational topics are probably better captured by other grouping concepts. This may include the philosophical and clinical orientation of BIPs, the training and skill sets of facilitators, and the characteristics and needs of client groups.

### **Frequency of Coverage of Educational Topics**

Reports of the frequency with which educational topics are covered by BIPs are generally aligned with their ratings of importance. In other words, the more important an educational topic was judged to help batterers end their abusive behavior, the more frequently that subject tended

to be explained or discussed in group sessions. For example, topics including accountability and taking responsibility for domestic abuse, denial and minimization of abuse, and time-out technique are correlated at a significance level of 5 percent or better, and reside within the list of top 10 educational topics in terms of importance and the frequency with which they are covered. (See Table 3-B for educational topics taught more frequently and Table 3-B1 of Appendix H for all other subjects.)

Exceptions to this rule include the topic of alcohol and substance abuse, where the rating of importance (22nd out of 53 potential subjects) was higher than the frequency with which it was covered (34th out of 53 potential subjects) in group. The importance rating of this topic may reflect the fact that many batterers participating in this study appear to be at risk for alcohol and substance abuse, while its moderate frequency of coverage may reflect caution by facilitators about focusing too much on topics that may provide their clients with excuses for their abusive behavior (e.g., I abused my spouse because I was drinking). Further, many program curricula are challenged with having to cover quite a number of important issues, and facilitators may believe that only fully developed substance abuse treatment programs can adequately help their clients. For these and other reasons facilitators might rate the subject of alcohol and substance abuse as quite important but cover it less frequently.

A second subject that is illustrative of the rare tendency for ratings of importance and frequency of coverage to diverge is clients using blame as a defense mechanism. In this case the relative frequency of coverage was quite high (4th out 53 possible subjects), while the rating of importance was somewhat lower (12th out of 53 subjects).



**Table 3-B. Frequency of Coverage of Educational Topics**

| Item | Educational Topics Explained or Discussed                                                             |                               |                          |                               | Average Frequency of Coverage, by Jurisdiction |      |      |      |      |      |
|------|-------------------------------------------------------------------------------------------------------|-------------------------------|--------------------------|-------------------------------|------------------------------------------------|------|------|------|------|------|
|      | Topics and Issues                                                                                     | Coding Cat.                   | N of BIPs Covering Topic | Average Frequency of Coverage | LA                                             | Riv  | SC   | SJ   | Sol  | Sig. |
| Q2A  | Accountability and taking responsibility for domestic abuse                                           | Accountability                | 43                       | 38.3                          | 34.0                                           | 46.5 | 46.5 | 37.8 | 33.4 | ns*  |
| Q43A | Power and control dynamics in abusive relationships                                                   | Power & Control               | 42                       | 36.4                          | 32.4                                           | 34.4 | 43.4 | 44.7 | 36.1 | ns*  |
| Q50A | Understanding the personal consequences of one's abusive behavior                                     | Abuse                         | 42                       | 33.8                          | 30.9                                           | 37.2 | 35.8 | 37.7 | 33.4 | ns*  |
| Q7A  | Blaming of others as defense mechanism                                                                | Defense Mechanisms: Batterers | 41                       | 32.7                          | 31.6                                           | 40.5 | 34.5 | 31.8 | 18.8 | ns   |
| Q17A | Denial of abuse as defense mechanism                                                                  | Defense Mechanisms: Batterers | 42                       | 32.2                          | 28.3                                           | 34.4 | 36.1 | 37.8 | 33.4 | ns   |
| Q35A | Minimization of abuse as defense mechanism                                                            | Defense Mechanisms: Batterers | 43                       | 31.9                          | 30.1                                           | 37.5 | 39.1 | 27.3 | 25.8 | ns   |
| Q51A | Violence prevention plan for client                                                                   | Planning                      | 37                       | 30.9                          | 30.3                                           | 30.7 | 27.4 | 39.5 | 33.4 | ns   |
| Q20A | Domestic abuse: What is it behaviorally? (e.g., emotional, economic, sexual, isolation, intimidation) | Abuse                         | 43                       | 30.5                          | 28.6                                           | 35.9 | 34.6 | 27.4 | 28.4 | ns   |
| Q49A | Time-out technique or procedure explained                                                             | Anger & Emotion Management    | 43                       | 30.2                          | 29.6                                           | 30.0 | 27.2 | 34.3 | 33.4 | ns   |
| Q29A | Identification of abuse triggers (anger, fear, grief, loss, separation, jealousy)                     | Anger & Emotion Management    | 42                       | 30.1                          | 27.1                                           | 36.1 | 24.0 | 37.8 | 33.4 | ns   |

Note 1: LA=Los Angeles, Riv=Riverside, SC=Santa Clara, SJ=San Joaquin, Sol=Solano

Note 2: ns=Nonsignificant differences among jurisdictions at the 5 percent level or beyond; ns\*=Nonsignificant differences with Tests for Homogeneity of Variance indicating that jurisdiction is not a reliable way to group BIPs for the corresponding item.

## Coping Skills Training

### Ratings of Importance: Coping Skills Training

Among facilitators responding to the PCS, training batterers in anger and emotion management emerged as among the most highly rated coping skills. More specifically, facilitators strongly endorsed the teaching of anger management and time-out techniques; they rated emotion expression skills training somewhat lower, though it still received high ratings of importance across jurisdictions. BIPs appear to be working solidly within the cognitive-behavioral school when they focus on the cognitive management and expression of emotion, for within this model poorly regulated emotion responses to stressful situations are thought to be important correlates of abuse among batterers.<sup>2</sup>

Ranking slightly below the most highly rated coping strategies in importance are conflict resolution skills, cognitive restructuring techniques to manage negative moods and self-talk, and positive forms of assertiveness training. Training clients to reflect and analyze their own behavior and life situations follows closely behind; clients learning to analyze their own behavior (to identify their abusive styles and areas of personal responsibility) and critical thinking skills were rated as important. (See Table 3-C for coping skills training rated of highest importance and Table 3-C1 of Appendix I for the full list.)

The high ratings of these cognitively oriented coping skills are consistent with anecdotal reports from clinicians and certainly with literature bearing on thinking and reasoning among batterers. All of this indicates that the decision to batter may be based on unexamined cognitive rules influencing batterer's coping responses<sup>3</sup> or misperceptions about what is actually at stake in an exchange with their domestic partners (e.g., their manhood, status as head of household, or their very survival).<sup>4</sup>

Of the remaining forms of coping skills training, all except three were rated somewhat important or higher across court jurisdictions. They include a mix of cognitive-behavioral techniques (e.g., alternative reactions to perceived problems, positive self-talk, countering techniques for irrational or problematic behavior, and thought switching and reframing), stress and coping training (e.g., relaxation and stress management training), interpersonal skills training (i.e., reflective listening training), and problem solving and planning (e.g., learning to manage one's finances and time). This rather broad approach to teaching coping skills suggests that BIPs may have taken a relatively integrative approach to intervention with batterers, focusing at various points during the intervention program on cognitive, emotional, behavioral, and social-interpersonal skills.

**Table 3-C. Importance of Coping Skills Training: Skills and Techniques**

| Item  | Coping Skills Training                                                                                                          |                                   |                          |                    | Average Rating of Importance, by Jurisdiction |     |     |     |     |      |
|-------|---------------------------------------------------------------------------------------------------------------------------------|-----------------------------------|--------------------------|--------------------|-----------------------------------------------|-----|-----|-----|-----|------|
|       | Skills and Techniques                                                                                                           | Coding Cat.                       | N of BIPs Covering Topic | Average Importance | LA                                            | Riv | SC  | SJ  | Sol | Sig. |
| Q1b.  | Anger management skills and techniques                                                                                          | Anger & Emotion Management        | 45                       | 4.6                | 4.5                                           | 5.0 | 4.3 | 4.8 | 5.0 | ns*  |
| Q21b. | Time-out technique training and practice                                                                                        | Anger & Emotion Management        | 44                       | 4.4                | 4.5                                           | 4.7 | 3.4 | 4.5 | 5.0 | ns*  |
| Q5b.  | Conflict resolution skills and/or techniques                                                                                    | Conflict Resolution & Negotiation | 45                       | 4.2                | 4.3                                           | 4.6 | 3.6 | 4.0 | 4.8 | ns*  |
| Q4b.  | Cognitive restructuring techniques to manage negative moods and negative self-talk                                              | Cognitive-Behavioral              | 45                       | 4.2                | 4.3                                           | 4.4 | 3.7 | 3.8 | 4.8 | ns*  |
| Q2b.  | Assertiveness training (while demonstrating respect for self and partner) as alternative to aggression                          | Interpersonal Skills              | 44                       | 4.0                | 4.3                                           | 4.0 | 3.1 | 3.7 | 4.5 | ns   |
| Q3b.  | Client practices analyzing his own behavior to identify the specifics of his abusive style and areas of personal responsibility | Cognitive-Behavioral (Duluth)     | 41                       | 3.9                | 3.9                                           | 3.4 | 4.4 | 3.7 | 4.3 | ns   |
| Q9b.  | Emotional expression skills training                                                                                            | Interpersonal Skills              | 44                       | 3.8                | 3.9                                           | 4.0 | 3.6 | 3.0 | 4.0 | ns   |
| Q7b.  | Critical thinking skills for clients/abusers                                                                                    | Cognitive-Behavioral (Duluth )    | 41                       | 3.7                | 4.1                                           | 3.4 | 3.7 | 2.5 | 4.3 | ns*  |
| Q14b. | Personal self-control techniques when parenting to avoid abusive behavior                                                       | Stress & Coping                   | 41                       | 3.7                | 3.8                                           | 4.3 | 3.0 | 3.7 | 3.3 | ns   |
| Q15b. | Positive self-talk training                                                                                                     | Cognitive-Behavioral              | 40                       | 3.7                | 3.9                                           | 3.9 | 3.1 | 2.8 | 4.8 | ns   |
| Q11b. | Alternative reactions to perceived problems or threats taught and practiced                                                     | Cognitive-Behavioral (Duluth)     | 41                       | 3.7                | 3.7                                           | 4.6 | 3.4 | 2.7 | 4.0 | ns   |
| Q19b. | Relaxation and stress management training                                                                                       | Stress & Coping                   | 41                       | 3.6                | 3.8                                           | 3.4 | 3.0 | 3.7 | 4.0 | ns   |
| Q13b. | Negotiation and compromise skills training                                                                                      | Conflict Resolution & Negotiation | 43                       | 3.5                | 3.7                                           | 3.7 | 3.6 | 2.7 | 3.5 | ns   |
| Q18b. | Reflective listening training                                                                                                   | Interpersonal Communication       | 41                       | 3.4                | 3.6                                           | 3.3 | 2.7 | 3.3 | 4.3 | ns   |
| Q6b.  | Countering technique for irrational or problematic beliefs                                                                      | Cognitive-Behavioral              | 37                       | 3.4                | 3.8                                           | 3.3 | 3.1 | 1.7 | 4.5 | ns   |
| Q16b. | Problem-solving skills training for dealing with everyday living, including managing finances and time                          | Problem Solving & Planning        | 39                       | 3.2                | 3.5                                           | 3.6 | 2.7 | 2.5 | 3.3 | ns   |
| Q10b. | Emotional sensitization exercises to help client learn to identify his emotions                                                 | Stress & Coping                   | 36                       | 3.1                | 3.5                                           | 3.1 | 2.1 | 2.5 | 4.0 | ns   |
| Q20b. | Thought switching and reframing training.                                                                                       | Cognitive-Behavioral              | 34                       | 3.0                | 3.2                                           | 3.7 | 1.7 | 1.8 | 4.5 | ns*  |

Note 1: LA=Los Angeles, Riv=Riverside, SC=Santa Clara, SJ=San Joaquin, Sol=Solano

Note 2: ns=Nonsignificant differences among jurisdictions at the 5 percent level or beyond; ns\*=Nonsignificant differences with Tests for Homogeneity of Variance indicating that jurisdiction is not a reliable way to group BIPs for the corresponding item.

The four forms of coping skills training that, on average, were thought to be of more marginal importance in abuse intervention included relatively focused cognitive-behavioral techniques typically employed in the treatment of individuals in a more formal therapeutic context.<sup>5</sup> They include decatastrophizing and depathologizing techniques, label shifting or relabeling training, and reattribution skills training. This last set of findings may be best understood when placed in the context of the preferred intervention mechanism employed by most BIPs in California.

The overwhelming majority of BIPs participating in this study work with court-ordered batterers in two-hour group sessions. In contrast, many of the techniques that are commonly used in traditional forms of cognitive-behavioral therapy were designed to be employed in individual sessions between a therapist and client. In these single-client sessions each program of treatment is tailored to the client's needs in strict accord with an extensive assessment process that highlights cognitive and behavioral strengths and weaknesses. This is not to suggest that cognitive-behavioral techniques cannot or are not effectively adapted to group treatment models. Rather it indicates that this is an inherently challenging process and may be less achievable for group facilitators who do not have access to specialized training or curriculum materials that fully support this approach.

### **Ratings of Importance: Coping Skills Training Versus Educational Topics**

A review of Tables 3-A and 3-C indicates that 19 educational topics were rated as ranging from important to very important in helping batterers end their abusive behavior. It is interesting to note that only 5 types of coping skills training attain this average level of importance. This may suggest that many BIPs responding to the PCS may place greater initial emphasis on the importance of helping clients understand their abuse and its implications, with somewhat less emphasis on training clients in new forms of coping with and adapting to stressors in their daily lives. This is consistent with certain abuse intervention models<sup>6</sup> as well as anecdotal evidence from interviews suggesting that some facilitators may try to tailor what is emphasized in their intervention programs to the developmental level of their clients. More specifically, as clients develop a deeper understanding of the causes and consequences of their domestic violence, are able to take greater responsibility for their abuse, and become more skilled in their coping behavior, facilitators may assign more advanced subjects and skills for them to learn.

### **Reliability of Court Jurisdiction as a Grouping Variable: Coping Skills Training**

Court jurisdiction again proved to be inconsistent in its reliability as a way to group BIPs in terms of their ratings of the importance of coping skills training. For example, court jurisdiction does not appear to be a statistically reliable way of classifying BIPs in terms of their importance ratings of the four most highly rated coping skills or of two of the four coping skills rated as least important (Table 3-C). This suggests that the approach BIPs are taking to training is probably influenced by factors beyond those including the court jurisdiction in which they reside, the local licensing requirements with which they must comply, and the justice system partners with whom they most frequently interact.

### **Frequency of Use: Coping Skills Training**

On average, the frequency with which BIPs responding to the survey teach coping skills is consistently related to their ratings of importance. For example, the two coping strategies rated highest are also most frequently taught by responding BIPs. (See Table 3-D for coping skills most frequently covered and Table 3-D1 of Appendix J for the full list.) In fact, the correlation between rating of importance of coping skill and frequency of use in group training is significant at the level of one-tenth of one percent or beyond ( $p < .001$ ) for all but one technique.

An exception to this rule is reattribution skills training, the cognitive-behavioral coping skill rated as of lowest importance across BIPs. The highly specific nature of this cognitive-behavioral technique, its association with formal approaches to cognitive-behavioral therapy, and its inconsistent use across BIPs may contribute to its divergence from the overall norm.

### **Ratings of Frequency of Use: Coping Skills Training Versus Educational Topics**

A review of Tables 3-B and 3-D indicates that 10 educational topics were reportedly used in 21 to 52 group sessions in the course of a 52-week intervention program (with data in these tables representing the midpoint of each frequency interval), while 2 types of coping skills training attained this intense level of use.

This pattern of findings parallels those described for ratings of importance of survey items by BIPs. It appears to add to the qualitative evidence suggesting that BIPs responding to the PCS place somewhat greater emphasis on helping their clients understand their abuse and its implications relative to training clients in new forms of coping and adaptation. However, these findings should not be interpreted as characterizing any single BIP in terms of its program emphasis, nor that the emphasis of programs remains the same throughout the course of a 52-week program. We may learn more through further analysis of this data, along with its triangulation with other independent sources of information.

**Table 3-D. Frequency of Coping Skills Training: Skills and Techniques**

| Coping Skills Training |                                                                                                                                 |                                   |                          |                               | Average Frequency of Coverage, by Jurisdiction |      |      |      |      |      |
|------------------------|---------------------------------------------------------------------------------------------------------------------------------|-----------------------------------|--------------------------|-------------------------------|------------------------------------------------|------|------|------|------|------|
| Item                   | Skills and Techniques                                                                                                           | Coding Cat.                       | N of BIPs Covering Topic | Average Frequency of Coverage | LA                                             | Riv  | SC   | SJ   | Sol  | Sig. |
| Q1b.                   | Anger management skills and techniques                                                                                          | Anger & Emotion Management        | 45                       | 31.9                          | 26.2                                           | 43.5 | 31.6 | 37.7 | 33.4 | ns   |
| Q21b.                  | Time-out technique training and practice                                                                                        | Anger & Emotion Management        | 44                       | 30.8                          | 29.6                                           | 38.9 | 27.5 | 30.7 | 28.1 | ns*  |
| Q5b.                   | Conflict resolution skills and/or techniques                                                                                    | Conflict Resolution & Negotiation | 45                       | 28.4                          | 27.0                                           | 34.5 | 27.2 | 27.3 | 28.1 | ns   |
| Q4b.                   | Cognitive restructuring techniques to manage negative moods and negative self-talk                                              | Cognitive-Behavioral              | 45                       | 26.3                          | 22.7                                           | 32.9 | 25.7 | 29.1 | 30.6 | ns   |
| Q2b.                   | Assertiveness training (while demonstrating respect for self and partner) as alternative to aggression                          | Interpersonal Skills              | 44                       | 21.2                          | 21.8                                           | 26.9 | 18.8 | 13.7 | 23.1 | ns   |
| Q3b.                   | Client practices analyzing his own behavior to identify the specifics of his abusive style and areas of personal responsibility | Cognitive-Behavioral (Duluth)     | 41                       | 29.7                          | 24.6                                           | 42.1 | 31.6 | 34.1 | 30.9 | ns   |
| Q9b.                   | Emotional expression skills training                                                                                            | Interpersonal Skills              | 44                       | 25.4                          | 24.2                                           | 30.1 | 25.6 | 23.6 | 25.6 | ns   |
| Q7b.                   | Critical thinking skills for clients/abusers                                                                                    | Cognitive-Behavioral (Duluth )    | 41                       | 24.9                          | 23.2                                           | 23.8 | 34.3 | 20.4 | 25.6 | ns   |
| Q14b.                  | Personal self-control techniques when parenting to avoid abusive behavior                                                       | Stress & Coping                   | 41                       | 20.9                          | 20.8                                           | 24.2 | 17.0 | 20.6 | 22.3 | ns   |
| Q15b.                  | Positive self-talk training                                                                                                     | Cognitive-Behavioral              | 40                       | 23.1                          | 23.5                                           | 22.6 | 20.5 | 25.5 | 23.1 | ns   |
| Q11b.                  | Alternative reactions to perceived problems or threats taught and practiced                                                     | Cognitive-Behavioral (Duluth)     | 41                       | 26.4                          | 23.9                                           | 35.9 | 34.2 | 17.5 | 20.6 | ns   |
| Q19b.                  | Relaxation and stress management training                                                                                       | Stress & Coping                   | 41                       | 20.6                          | 20.2                                           | 27.3 | 13.8 | 25.8 | 18.0 | ns   |

Note 1: LA=Los Angeles, Riv=Riverside, SC=Santa Clara, SJ=San Joaquin, Sol=Solano

Note 2: ns=Nonsignificant differences among jurisdictions at the 5 percent level or beyond; ns\*=Nonsignificant differences with Tests for Homogeneity of Variance indicating that jurisdiction is not a reliable way to group BIPs for the corresponding item.

## Teaching Strategies and Techniques

### Ratings of Importance: Teaching Strategies and Techniques

Techniques that give the group facilitator a central role in the teaching and training in groups, techniques that emphasize insight into and accountability for one's abuse, and rehearsal of new forms of positive coping behavior were rated as more important by BIPs responding to the survey. More specifically, group facilitators rated the following teaching techniques as ranging from very important to moderately important. (See Table 3-E for teaching strategies and techniques rated of highest importance and Table 3-E1 of Appendix K for the full list.).

- Group discussions structured and led by a facilitator
- Analysis by clients of their own abusive behavior and anger triggers
- Facilitator's therapeutic/educational confrontation of clients
- Challenging attitudes and beliefs that encourage abuse by group members
- Facilitator's leading clients through a description of some to their most severe incidents of abuse
- Lectures or formal presentations by facilitator
- Homework focused on clients' plans for ending their abuse
- Role-playing led by the facilitator
- Rehearsal of cognitive-behavioral strategies in group
- Rehearsal of coping strategies (e.g., time-out technique)

The influences of both the Duluth and cognitive-behavioral approaches to instruction are clearly evident in many of these teaching strategies and techniques. In any given BIP, however, the approaches may well be configured to represent a hybrid approach to batterer intervention as BIPs are taking a view to teaching that draws upon a number of the most prominent models in domestic violence intervention.

Instructional techniques that were rated as of clearly lower importance included those that emphasize attachment issues and strategies for addressing them (e.g., female facilitators lead groups to address gender-based issues of client trust); the use of advanced students as discussion or role-play leaders; employing quizzes and tests to check on clients' learning and mastery of course content; and various forms of homework requiring reading and writing assignments that focus on some aspect of a batterer's abusive attitudes, beliefs, or behavior.

### **Ratings of Frequency of Use: Teaching Strategies and Techniques**

The ratings for the frequency of use of educational strategies and techniques parallel the ratings of their importance (see Table 3-E for teaching strategies and techniques used most frequently and Table 3-E1 of Appendix K for the full list), which is to suggest that the more important a technique was rated, the more frequently it was used. In fact, correlations between importance and frequency are quite substantial, never dropping below a zero order correlation of .6 and all significant at the level of one-tenth of one percent or beyond ( $p < .001$ ).

### **Jurisdictional Differences: Teaching Strategies and Techniques**

Only a single jurisdictional difference in the ratings of importance of instructional strategies was found—the use of films and videos not specifically made for domestic violence courses but relevant to domestic abuse. This suggests again that there is considerable consistency across

BIPs in the way they approach teaching batterers. The form of this cross-jurisdictional agreement appears to place the group facilitator at the center of instruction, teaching clients to become accountable for and critically aware of their abuse while emphasizing positive forms of thinking and behaving as alternatives to abusive behavior (See Table 3-E and Table 3-E1 of Appendix K).

### Reliability of Court Jurisdiction as a Grouping Variable: Teaching Strategies and Techniques

As noted before, the jurisdiction within which BIPs are located has no measurable effect on BIPs in relation to the ratings of the importance of teaching strategies and the frequency with which they are used in group (Table 3-E and Table 3-E1 of Appendix K).

**Table 3-E. Importance and Frequency of Teaching Strategies and Techniques**

| Teaching Strategies and Techniques |                                                                                                                                                                       |                          |         |       | Average Rating of Importance and Frequency, by Jurisdiction |     |     |     |     |      |
|------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|---------|-------|-------------------------------------------------------------|-----|-----|-----|-----|------|
| Item                               | Strategies and Techniques                                                                                                                                             | N of BIPs Covering Topic | Average |       | LA                                                          | Riv | SC  | SJ  | Sol | Sig. |
|                                    |                                                                                                                                                                       |                          | Import. | Freq. |                                                             |     |     |     |     |      |
| 7c.                                | Group discussion: Structured and led by facilitator                                                                                                                   | 44                       | 4.6     | 41.9  | 4.4                                                         | 5.0 | 4.6 | 4.7 | 5.0 | ns*  |
| 1c.                                | Client instructed in the analysis of his own abusive behavior to become aware of personal anger triggers and other aspects of his abusive style and cycle of violence | 43                       | 4.3     | 30.5  | 4.3                                                         | 4.4 | 3.6 | 4.8 | 5.0 | ns*  |
| 26c.                               | Therapeutic/educational confrontation of clients by group facilitator                                                                                                 | 41                       | 4.0     | 33    | 4.0                                                         | 3.7 | 3.9 | 3.8 | 4.8 | ns   |
| 9c.                                | Group members allowed to take the lead in challenging attitudes and beliefs that encourage domestic violence                                                          | 40                       | 3.7     | 29.8  | 3.6                                                         | 2.9 | 4.7 | 3.3 | 4.0 | ns*  |
| 3c.                                | Facilitator leads client through a description of some of his most severe incidents of partner abuse                                                                  | 40                       | 3.6     | 23.6  | 3.5                                                         | 3.1 | 4.3 | 3.7 | 3.8 | ns   |
| 18c.                               | Lecture or formal presentation by facilitator                                                                                                                         | 37                       | 3.5     | 31.6  | 3.4                                                         | 3.7 | 3.7 | 3.5 | 3.3 | ns   |
| 10c.                               | Homework: Client develops prevention or safety plan to prevent future abuse                                                                                           | 38                       | 3.4     | 21.3  | 3.2                                                         | 3.0 | 3.7 | 3.7 | 4.3 | ns   |
| 24c.                               | Role-playing led by group facilitator                                                                                                                                 | 39                       | 3.3     | 17.5  | 3.8                                                         | 2.7 | 3.0 | 2.2 | 4.3 | ns   |
| 21c.                               | Rehearsal of cognitive and behavioral skills in group                                                                                                                 | 37                       | 3.3     | 24    | 3.5                                                         | 3.1 | 3.4 | 2.3 | 3.8 | ns   |
| 22c.                               | Rehearsal of coping strategies (e.g. time-out).                                                                                                                       | 35                       | 3.3     | 26.2  | 3.6                                                         | 3.3 | 2.7 | 2.0 | 4.5 | ns   |
| 25c.                               | Therapeutic/educational confrontation of clients by "advanced students/clients" in group sessions                                                                     | 33                       | 3.1     | 25.3  | 3.5                                                         | 3.0 | 2.9 | 2.2 | 2.8 | ns   |
| 6c.                                | Films and videos: Developed specifically for domestic violence courses                                                                                                | 35                       | 2.9     | 11.5  | 2.9                                                         | 3.1 | 2.9 | 2.0 | 4.3 | ns   |

Note 1: LA=Los Angeles, Riv=Riverside, SC=Santa Clara, SJ=San Joaquin, Sol=Solano

Note 2: ns=Nonsignificant differences among jurisdictions at the  $p < .05$  level or beyond; ns\*=Nonsignificant differences with Tests for Homogeneity of Variance indicating that jurisdiction is not a reliable way to group BIPs for the corresponding item.

### Summary of Findings

The educational subjects identified as important by BIPs in helping batterers end their domestic abuse appear consistent with legislative mandates intended to hold offenders accountable for



their abusive behavior. The program elements that were rated more highly in importance also appear to be central to some of the most influential domestic violence intervention models developed in this county. These include efforts to (1) hold batterers accountable and personally responsible for their domestic violence; (2) make batterers aware of (and change) the attitudes and beliefs that underpin their abusive behavior, including issues related to power and control as well as the management of anger, emotion, and stress in domestic situations; and (3) give batterers an understanding of the effects and implications of domestic abuse.

It is important to note that no significant differences were found among jurisdictions in terms of the average ratings with which facilitators assigned importance to educational topics rated as important to very important for helping batterers. However, statistically significant differences were noted among a number of educational topics thought by respondents to be less important in ending domestic abuse. Variation in the ratings of the importance of these topics as well as the highly correlated frequency with which they are taught may be influenced by facilitators' judgments concerning the needs of the client groups they serve, the beliefs prevalent in BIPs about the etiology of domestic violence, and different interpretations among BIPs regarding the requirements of local and state mandates for the treatment of domestic abuse.

Anger and emotion management emerged among facilitators as some of the most highly rated coping skills for batterers to learn. The high ratings of these cognitively oriented coping skills are consistent with the view that domestic violence is often based on the batterers' misperceptions that their domestic status and even image of themselves as men are at stake in contentious domestic situations and must be defended at all costs. Of the remaining forms of coping skills training, the majority were rated as somewhat important or higher across court jurisdictions. They include a mix of cognitive-behavioral techniques, stress and coping training, interpersonal skills training, and problem solving and planning. This rather broad approach to teaching coping skills suggests that BIPs have taken a relatively integrative approach to intervention with batterers, focusing at various points during the intervention program on cognitive-behavioral, emotional, and social-interpersonal skills.

Many BIPs responding to the PCS appear to place greater emphasis on the importance of helping clients understand their abuse and its implications than they place on training clients in new forms of coping with and adapting to stressors in their domestic lives. This approach is consistent with a number of prominent abuse intervention models,<sup>7</sup> as well as anecdotal evidence from interviews suggesting that intervention program directors may believe that they initially need to focus on helping batterers develop a basic understanding of their abuse, its proximal origin and implications, and a few simple coping strategies. They may introduce more demanding forms of coping skills once they've developed a foundation of basic awareness. This in turn suggests that some facilitators tailor the curriculum to the educational and developmental levels of their clients.

The frequency with which BIPs report teaching educational topics appears to vary with their ratings of its importance. An even stronger relationship was observed between ratings of importance and the frequency of coverage for coping skills training. In other words, the more important an educational topic or coping skill is judged to be by facilitators, the more frequently

it appears to be addressed in group. Anecdotal evidence from interviews with highly trained senior facilitators suggests that there may also be important nuance in this approach. Some important topics may be introduced at critical points when batterers are developmentally ready while other, less important topics, are covered to facilitate the development and readiness of clients.

Approaches to teaching that emphasize insight and accountability by batterers into their abusive behavior, approaches that involve facilitators centrally in the teaching and training that occurs in group, and strategies that emphasize the rehearsal of new forms of positive coping behavior were rated as more highly important by BIPs responding to the survey. It was noted earlier that the influences of both cognitive-behavioral and Duluth approaches were evident in many of these teaching strategies and techniques, they may be configured by facilitators into hybrid approaches to intervention that reflect their own training and background as well as perceptions of the needs of their client groups. Ratings of the frequency of use of teaching strategies and techniques assessed by the PCS appear to underline this last fact, with BIPs more frequently using those approaches they rated as more useful in helping clients end their abuse.

Court jurisdiction was generally not useful for grouping BIPs in relation to their approaches to training batterers in new forms of positive coping or the approaches to teaching that they employ in the treatment of batterers. Further, no statistically significant differences were observed in importance ratings of the four most highly rated coping skills or in two of the four coping skills rated as least important. This suggests that the approach BIPs are taking to teaching and training may be influenced by system-level factors beyond those of jurisdiction.

## **Summary**

In this chapter we discuss how the Program Content Survey was developed to try to better understand the content and educational methods used in 52-week domestic violence programs. The findings indicate that program curriculums are consistent with legislative mandates. However, what little differentiation exists between programs as measured on the Program Content Survey does not provide enough evidence to determine whether certain methods used in BIPs yield better outcomes. Greater differentiation was found among offender characteristics, however, which we examine in the next chapter.

## Endnotes Chapter 3

---

1. D. Dutton and D. Sonkin, eds., *Intimate Violence: Contemporary Treatment Innovations* (Binghamton, NY: The Haworth Press, Inc. 2002).
2. L. K. Hamberger and J. M. Lohr, "Proximal Causes of Spouse Abuse: A Theoretical Analysis for Cognitive-Behavioral Interventions," in P. L. Caesar and L. K. Hamberger, eds., *Treating Men Who Batter: Theory, Practice and Programs* (New York: Springer, 1989); E. Pence and M. Paymar, *Education Groups for Men Who Batter: The Duluth Model* (New York: Springer, 1993).
3. C. Murphy and C. Eckhardt, *Treating the Abusive Partner: An Individualized Cognitive-Behavioral Approach* (New York: The Guilford Press, 2005).
4. E. Pence and M. Paymar, *id.* note 2.
5. H. Sinclair, "A Community Activist Response to Intimate Partner Violence," in E. Aldarondo and F. Mederos, eds., *Programs for Men Who Batter: Intervention and Prevention Strategies in a Diverse Society* (Civic Research Institute, 2002).
6. *Id.* note 3.
7. *Id.* note 5.

## Chapter 4: Offender Profiles

### ***Introduction***

Gathering detailed information on offender characteristics is a critical component of this study. The purpose, however, is not so much to help increase our understanding of domestic violence offenders, such as which characteristics contributed to their abusive behavior or which set of risk or protective factors are associated with their different propensities for compliance. Rather, the need for offender profile data arises from the non-experimental nature of the study design, in which study subjects recruited from different jurisdictions may exhibit different characteristics, and these different characteristics may lead to different propensities for compliance independent of any system-level impacts that might exist. With system-level impacts as the primary focus of this study, offender profiles thus provide a means of rendering statistically more comparable the study samples across the different jurisdictions. In other words, they function as control variables in a multivariate analysis framework, which is discussed in detail in Chapter 5.

A brief descriptive analysis of offender profiles, however, will provide essential context leading to the analysis of outcome measures in the following chapter. In Table 4-A measures of offender characteristics are grouped into four categories: (1) family relations, including relationship with the victim (wife or girlfriend) and children, and living arrangements with them; (2) socioeconomic status, including income, employment, education, and race/ethnicity; (3) measures of criminal history constructed from the California State Department of Justice (DOJ) arrest records and CAGE indicator of alcohol/drug abuse; and (4) indicators of abusive behavior and conflicts with the victim, as measured by the revised Conflict Tactics Scale 2 (CTS2) concerning the frequency of various forms of conflict (as well as positive, non-abusive interactions with the victim) in the past year. Measures that vary across the jurisdictions at statistically significant levels (1 and 5 percent levels) are indicated in the table.

### ***Family Relations***

There are noticeable differences across the jurisdictions regarding the relationship between the offender and his victim and children, as well as in his living arrangements with them. Overall, approximately 40 percent of the offenders were living with the victim at the time of program enrollment. By jurisdiction, it varies from a low of 35 percent in San Joaquin County to a high of 50 percent in Solano County. With regard to relationship with the victim, slightly less than one-half (45 percent) involved either current or former wife, with no statistically significant difference across the jurisdictions.

Significant differences exist across the jurisdictions in the proportion of offenders who had children *and* were living with them at the time of program enrollment, ranging from 28 percent in Los Angeles County to 48 percent in Solano County.

**Table 4-A . Offender Characteristics, by Jurisdiction**

| Risk Factors                                           | Los Angeles | Riverside | Santa Clara | Solano   | San Joaquin | Total Sample | Valid Sample Size |
|--------------------------------------------------------|-------------|-----------|-------------|----------|-------------|--------------|-------------------|
| <b>Family Relations</b>                                |             |           |             |          |             |              |                   |
| Percent Living with Victim**                           | 36%         | 46%       | 36%         | 50%      | 35%         | 38%          | 1,405             |
| Percent Victim Was Wife (current or former)            | 47%         | 42%       | 42%         | 48%      | 44%         | 45%          | 1,411             |
| Percent Living with Children**                         | 28%         | 40%       | 32%         | 48%      | 36%         | 33%          | 1,384             |
| <b>Socio-economic Status</b>                           |             |           |             |          |             |              |                   |
| Percent with Some College**                            | 27%         | 19%       | 27%         | 25%      | 13%         | 23%          | 1,376             |
| Percent Employed Full-Time                             | 50%         | 53%       | 45%         | 51%      | 42%         | 47%          | 1,176             |
| Percent Lost Job in Past Year                          | 21%         | 20%       | 24%         | 17%      | 23%         | 22%          | 1,167             |
| Average Annual Income**                                | \$17,324    | \$20,865  | \$20,086    | \$16,113 | \$10,976    | \$17,489     | 1,146             |
| Percent African American**                             | 20%         | 13%       | 9%          | 32%      | 18%         | 17%          | 1,361             |
| Percent Hispanic**                                     | 58%         | 45%       | 57%         | 15%      | 46%         | 51%          | 1,361             |
| Percent White**                                        | 12%         | 34%       | 22%         | 28%      | 27%         | 22%          | 1,361             |
| Percent "Other"***                                     | 9%          | 8%        | 11%         | 25%      | 9%          | 10%          | 1,361             |
| Percent Needing Interpreter**                          | 34%         | 12%       | 21%         | 7%       | 10%         | 22%          | 1,457             |
| Average Age at Intake                                  | 33.90       | 33.77     | 33.93       | 34.85    | 32.66       | 33.68        | 1,328             |
| <b>Criminal History and Drug/Alcohol Abuse</b>         |             |           |             |          |             |              |                   |
| Average Age at First Arrest**                          | 25.87       | 24.69     | 24.90       | 24.30    | 22.90       | 24.78        | 1,301             |
| Average Number of Prior Arrests for All Offenses**     | 5.56        | 6.48      | 7.70        | 7.49     | 7.90        | 6.87         | 1,303             |
| Average Number of Prior Arrests for Assault Offenses** | 2.43        | 2.95      | 3.05        | 3.76     | 3.65        | 2.99         | 1,303             |
| Average Number of Prior Arrests for DV Offenses**      | 1.55        | 1.91      | 2.11        | 2.36     | 2.36        | 1.96         | 1,303             |
| Average Number of Prior Arrests for Drug Offenses**    | 1.27        | 1.63      | 2.48        | 2.00     | 1.96        | 1.85         | 1,303             |
| Average Number of Prior Arrests for Felony Offenses    | 3.25        | 3.87      | 3.21        | 3.67     | 3.91        | 3.46         | 1,303             |
| Average CAGE Score (0-4)                               | 1.15        | 1.27      | 1.42        | 1.25     | 1.24        | 1.26         | 1,164             |
| <b>Abusive Behavior Indicators (CTS2)</b>              |             |           |             |          |             |              |                   |
| Negotiation**                                          | 57.59       | 62.77     | 66.07       | 71.75    | 57.97       | 61.53        | 1,237             |
| Psychological Aggression                               | 22.15       | 27.46     | 24.85       | 30.34    | 23.56       | 24.35        | 1,237             |
| Physical Aggression                                    | 7.24        | 7.78      | 5.91        | 7.63     | 7.00        | 6.93         | 1,237             |
| Injury (of Offender) Resulting from Conflicts          | 2.77        | 2.06      | 2.18        | 3.05     | 2.02        | 2.39         | 1,237             |
| Sexual Coercion                                        | 3.31        | 3.36      | 1.85        | 2.71     | 2.82        | 2.79         | 1,237             |

\*\* Differences across jurisdictions statistically significant at 1 percent level; \* significant at 5 percent level.

Further analysis reveals not only correlations among the three family-relation variables—whether the offender was living with the victim, was married to the victim, or was living with children—but also different subgroup patterns across the jurisdictions. With each of the three binary variables representing two subgroups, a total of eight subgroups can be created when the three variables are combined. The distribution of these smaller subgroups shows that about three-quarters of the total sample fall into three major categories:

- The largest subgroup consists of offenders whose victims are other than their wives, and who were not living with their victims or with any children at the time of program enrollment. This subgroup represents approximately 35 percent of the total sample.
- The second major subgroup is made up of offenders whose victims are their wives. However, they were not cohabitating at the time of program enrollment, and no children were staying with them either. Approximately 20 percent of the total belongs to this subgroup.
- The third major subgroup consists of offenders whose victims are their wives and who were living with their victims and their children. Slightly less than 20 percent of the total falls into this subgroup.

Our data further suggests the existence of correlations between each of the three variables and other socioeconomic characteristics of the offenders. Belonging to one rather than another of the three paired subgroups—living with the victim, being married to the victim, or living with children—appears to contribute to a positive (or negative) correlation with employment status and income. Thus, an offender whose wife is the victim, who is living with the victim, and who has children living with him as well is more likely to be employed and earning a higher income. Offenders with these “positive” characteristics are also shown to have a less extensive criminal arrest record; they also tend to be older. Level of education completed, however, is found not to be correlated with family-relation characteristics.

**Table 4-B. Major Subgroups Based on Offender’s Relationship with Victim and Children**

|                                                            | Los Angeles | Riverside | Santa Clara | Solano | San Joaquin | Total | Valid Sample Size |
|------------------------------------------------------------|-------------|-----------|-------------|--------|-------------|-------|-------------------|
| Victim other than wife, not living with victim or children | 35%         | 33%       | 38%         | 30%    | 36%         | 36%   | 485               |
| Victim is wife, not living with victim or children         | 23%         | 13%       | 20%         | 8%     | 19%         | 19%   | 263               |
| Victim is wife, living with victim and children            | 16%         | 22%       | 17%         | 26%    | 17%         | 18%   | 242               |
| Subgroup Total                                             | 74%         | 69%       | 74%         | 64%    | 72%         | 73%   | 990               |

Whether viewed theoretically as stake-in-conformity indicators (investment in status quo such as marriage and employment that may act as deterrence from non-conforming behaviors) or as measures of the degree to which the offender’s life was in a settled or unsettled condition (associated with more or less stress), the significant differences across the jurisdictions in offender profiles, as shown in Table 4-A, suggest that these are important “control” variables that need to be addressed in comparing outcomes across the jurisdictions.<sup>1</sup>

For the largest subgroup—those who lack any of the “positive” elements as measured by the family-relation variables—the data shows that Santa Clara County has the largest proportion of these cases at 38 percent of total within the jurisdiction, and Solano County has the smallest proportion at 30 percent.

The study sample in Solano, again, consists of the lowest proportion (8 percent) of the second subgroup (wife being the victim, but not living with the victim or children) among the jurisdictions. Riverside County has a similarly smaller proportion of these cases at 13 percent of total, relative to approximately 20 percent in the three remaining jurisdictions.

Consistent with the patterns for the first two subgroups, the data further shows that the Solano sample contains a higher proportion (26 percent) of the third subgroup than the other jurisdictions. As noted above, those in the third subgroup reveal more of what appear to be “positive” characteristics, suggesting perhaps that the offenders are in a relatively more settled situation in their life. Riverside County shows a slightly lower proportion of these cases (22 percent) relative to Solano, but noticeably higher than the other three jurisdictions.

Regardless of the specific mechanisms by which family dynamics may affect the offender’s performance in terms of either program compliance or re-offense behavior, the different subgroup compositions described above point to significantly different offender profiles across the jurisdictions, which need to be taken into consideration in outcome analysis.

Other than the three large subgroups described above, each of the remaining five subgroups—e.g., offenders who lived with a victim who was not their wife but lived with children or offenders who live with a victim who is their wife but does not live with children—represents a small fraction of the total sample, ranging from 3 to 9 percent.

## ***Socioeconomic Status***

Various socioeconomic indicators in Table 4-A depict an offender population that is generally of low status, with limited education, a low employment rate, and low income—and overwhelmingly of minority race/ethnicity background.

Overall, more than one-quarter of the total sample attained a high school education or less, while merely 5 percent graduated from college. Educational levels are lowest in San Joaquin County, with less than 15 percent having attended some college and less than 2 percent finishing an undergraduate or higher degree.

While the data shows statistically significant differences in educational attainment across the jurisdictions, only marginally significant differences exist in employment status. Overall, slightly less than one-half (47 percent) of the sample were fully employed at the time of program enrollment, varying from a low of 42 percent in San Joaquin County to a high of 53 percent in Riverside County. The data further indicates about 20 percent experienced job loss or work hour

reduction in the past year, and only approximately 40 percent of the total sample had stable employment in the recent past.

Further analysis of the relationship between educational attainment and employment status shows that only a small fraction of those with college degrees were faring better in employment, with approximately 60 percent fully employed—higher than the rest of the sample by a little more than 10 percentage points. For those without a college degree, increments of additional education—from “less than high school” to “high school graduate” to “some college, including associate’s degree”—do not seem to have had any impact on the employment status of the offenders.

The lack of a strong correlation between education and employment, which is typically expected to exist in the general population, as well as the overall low-level, marginal employment the study subjects, suggests the existence of some underlying, persistent, unobserved factors influencing both education and employment status of the offenders, for which the various socioeconomic indicators are serving as close proxies.

Self-reported income provides further evidence of the offenders’ low socioeconomic status consistent with the results from other variables discussed above. With nearly 20 percent of the sample reporting no income in the past one year, the overall average income reported is less than \$20,000 per year (an average of \$17,500 and a median of \$14,500). The study subjects in San Joaquin include an especially high percentage reporting no income at all—36 percent of the total relative to approximately 20 percent in Solano County and 15 percent in the other three jurisdictions.

When subgroups defined by various income and educational levels are examined in conjunction, the data shows that nearly 60 percent of the total sample had no more than a high school education along with a reported annual income of less than \$25,000. At the high end of the income-education continuum, a mere 2 percent of the total sample consists of college graduates with a reported income higher than \$40,000 per year.

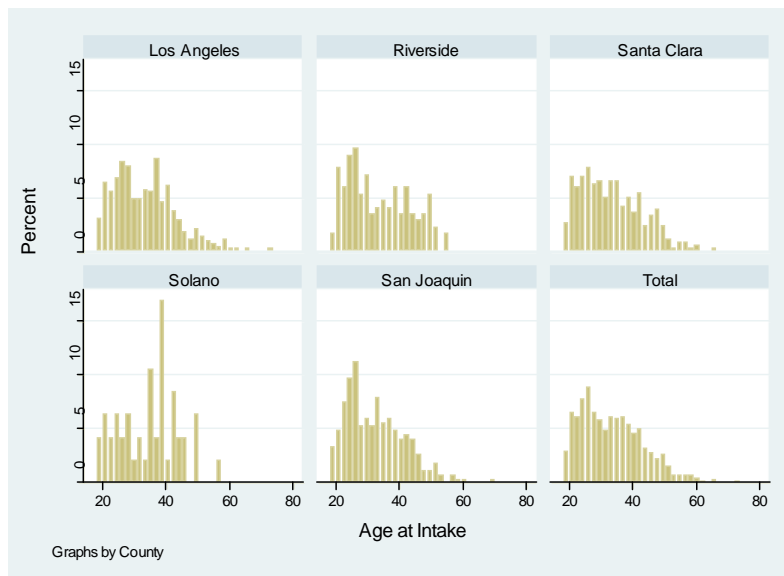
The racial/ethnic composition of the sample provides yet another important indicator regarding the status of the offenders, with nearly 80 percent of the total coming from minority groups. As Table 4-A shows, Hispanics are the largest group overall at approximately 50 percent of the total, followed by 22 percent whites, 17 percent African Americans, and 10 percent “others.” Furthermore, approximately 35 percent of Hispanics and “others”—or 22 percent of the entire sample—speak a native language other than English, indicating their status as recent immigrants.

Differences in the racial/ethnic composition of the samples are quite substantial across the jurisdictions. In Los Angeles and Santa Clara Counties, Hispanics account for almost 60 percent of the total, whereas they represent merely 15 percent of the total in Solano County. Los Angeles has the lowest proportion of whites at 12 percent, compared with 34 percent in Riverside County. African Americans are the smallest group in Santa Clara (9 percent of the total) but the largest group in Solano (32 percent of the total).

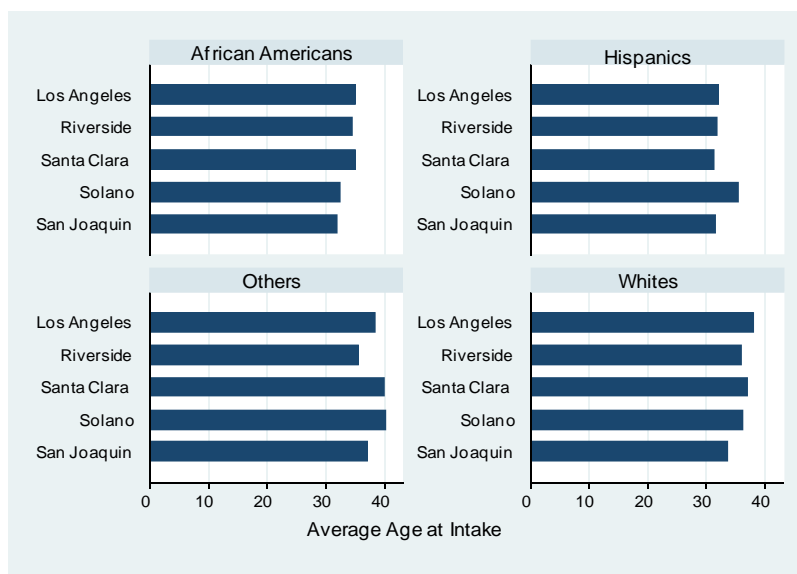


At first glance, the offender's age at intake appears to be fairly comparable across the jurisdictions, at an average of approximately 34 years. Ranging from 18 to 74, the overall age distribution is skewed to the right—the higher end of the continuum, as shown in Figure 4-A, with slightly over 40 percent under the age of 30. About one-third of the total sample are between the ages of 30 and 40, with the remaining 25 percent over the age of 40.

**Figure 4-A. Distribution of Age at Intake, by Jurisdiction**



**Figure 4-B. Average Age at Intake, by Jurisdiction and Race/Ethnicity**

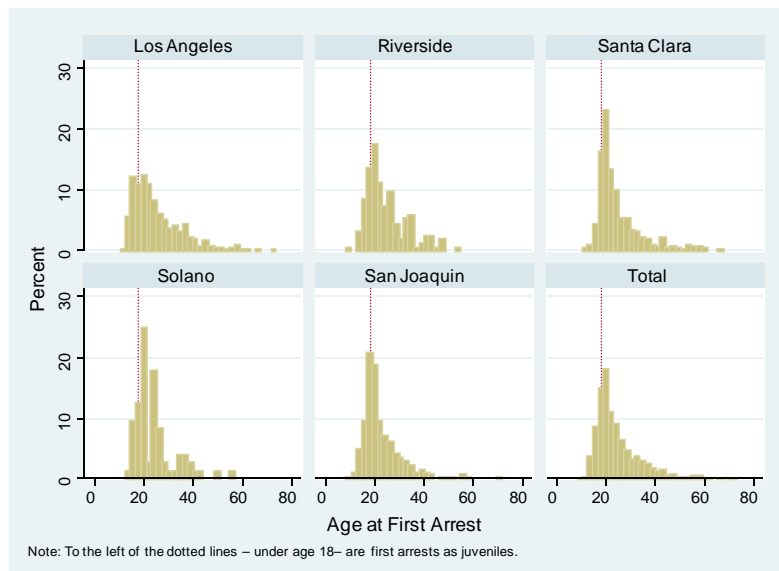


Further analysis reveals that age at intake varies significantly by race/ethnicity groups. Hispanics tend to be younger, with an average age of 32; they are followed by African Americans at 34, whites at 36, and finally “others” at a little over 38. Within each race/ethnicity group, there is also evidence of some variance across the jurisdictions, as shown in Figure 4-B. For example, the average age of African Americans is approximately 32 in San Joaquin, compared with 35 in Los Angeles. There is a similar difference between San Joaquin and Los Angeles for whites, with averages of approximately 34 and 38, respectively. When differences across jurisdictions in racial/ethnic composition are taken into consideration, it is clear that offenders in San Joaquin tend to be younger as a whole, whereas differences in other jurisdictions can be attributed to their different race/ethnicity.

## ***Criminal History and Drug/Alcohol Abuse***

Age at first arrest serves as an indicator of an offender’s age for onset of criminal activities. Combined with age at intake, this variable also provides information regarding the length of time that an offender has been engaged in criminal activities.

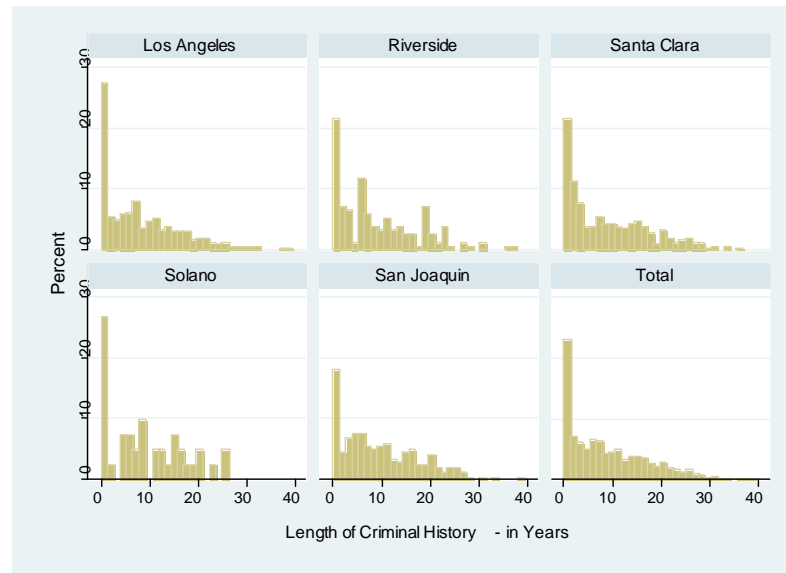
**Figure 4-C. Distribution of Age at First Arrest, by Jurisdiction**



From both perspectives—age of onset and length of criminal history—the data reveals some noticeably different patterns across the jurisdictions, as shown in Figure 4-C and Figure 4-D. An analysis of distributional patterns of age at first arrest, as shown in Figure 4-C, indicates that approximately 17 percent of the total sample were first arrested as juveniles (under the age of 18). These juvenile-onset offenders constitute the smallest proportion of the total in Santa Clara County at 9 percent of the total, compared with approximately 15 percent in Riverside and Solano Counties, and slightly over 20 percent in Los Angeles and San Joaquin Counties. Different distributional patterns also exist across the jurisdictions for those whose first arrest occurred over the age of 18—the late-onset offenders—resulting in an overall average age of

approximately 26 in Los Angeles at the highest end, compared with 23 in San Joaquin at the lowest end.

**Figure 4-D. Length of Criminal History, by Jurisdiction**



Just as differences in offender age at intake across the jurisdictions partly reflect the different racial/ethnic composition of the sample, it is important also to consider race/ethnicity in examining the age at first arrest. African Americans tended to have the first arrest on their criminal record at a younger age than other groups. Using African Americans as the comparison group, Hispanic offenders experienced their first arrest 2 years later, whites 4 years later, and “others” 10 years later than African Americans in the sample.

A complicating factor in the analysis above is related to offenders who recently moved to the country, which may result in incomplete, truncated arrest records available in the DOJ arrest database. If we control for immigrant status by using the flag in the court record that indicates “interpreter needed” as a proxy for recent immigration, then there is no longer any difference between Hispanics and African Americans in age of first arrest. In other words, Hispanics who did not need an interpreter, suggesting that they have been in the country longer than those who did need an interpreter, have a first arrest on their criminal record at about the same age as African Americans in the sample.

The later onset age for “others” is also reduced from 10 to 8 years, again relative to African Americans. Compared within Hispanics and “others,” two subgroups with significant numbers of non-English speakers, the age at first arrest for non-English-speaking Hispanics is approximately 6 years older than for their English-speaking counterparts (28 versus 22), and for non-English-speaking “others” 11 years (39 versus 28). While truncated arrest records could be partly responsible for the differences noted above, it is also reasonable to assume that non-English

speakers are differentiated from their native counterparts by other unobserved characteristics as well, resulting in real differences in criminal history records.

Overall, the length of criminal history for offenders in the sample is approximately 9 years (median of 7 years) for the entire sample, with noticeable differences across the jurisdictions, as shown in Figure 4-D. In all jurisdictions, a substantial proportion of offenders had been in the criminal justice system for a relatively short time, less than one to two years from first arrest to intake. Overall they represent approximately 25 percent of the total, ranging from approximately 20 percent in Solano and San Joaquin to more than 30 percent in Los Angeles (the sum of the first two bars on the left side of each county's histogram). While the offenders in Los Angeles appear to have a shorter criminal history, with offenders in Solano and San Joaquin on the high end, the appearance of the differences again can be attributed to the composition of different race/ethnicity groups and non-English speakers.

Closely correlated to the length of criminal history are the various frequency measures of prior arrests shown in Table 4-A, including counts of all offense types, assaults, domestic violence offenses, drug- and alcohol-related offenses, and felonies. The overall average number of prior arrests including all offense types is close to 7 (median of 5), with significant variances across the jurisdictions. The average arrest counts range from a high of 7.9 in San Joaquin to a low of 5.6 in Los Angeles.

Further breakdown by offense type shows that, on average, approximately half of the prior arrest records consist of assault charges (overall average of 3); approximately 40 percent involving domestic violence charges (overall average of 2, as a subset of the general assault category); and approximately 20 percent related to drug and alcohol charges. The relatively high frequency of prior arrests, along with the diversity of offense types, suggests a profile of chronic offenders whose recent domestic violence offense could be merely an episode in the trajectory of a long criminal career.

Analysis of the full range of prior arrest records (from 1 to 48) also reveals that, despite an overall profile exhibiting characteristics of chronic offenders, for approximately 20 percent of the total sample the offense that led to their current conviction was their first arrest on record of any offense type. Consistent with the analysis above on average prior arrests, offenders with a single prior arrest constitute the largest proportion in Los Angeles at approximately 25 percent of the total, compared with 18 percent in Riverside and approximately 14 percent in the other three jurisdictions.

It is interesting to note that, while it is reasonable to expect older age to be associated with longer arrest records, there is no significant difference across the age groups in the proportion of first offenders. For those under the age of 25 at intake, the proportion of first offenders is slightly higher at 22 percent of the total; for the other age groups of 25 to 30, 30 to 40, and over 40, the relevant percentages are 17, 17, and 19, respectively.

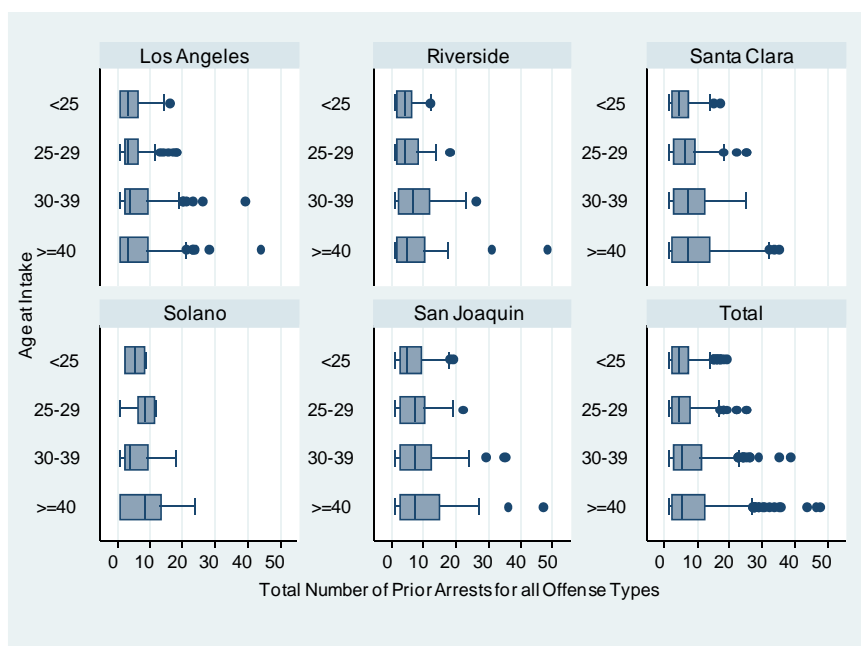
The different patterns of prior arrest records by age groups can be seen most revealingly by box-whisker plots, as shown in Figure 4-E, with different age subgroups shown separately within each jurisdiction. With the horizontal axis representing the number of prior arrests, the box width

represents the dispersion of cases between the first and third quartiles (25th and 75th percentiles) of the samples with respect to their prior arrests; the vertical line within each box represents the median value of prior arrests within each subgroup. The whiskers, stretching out in both directions from the median value, along with dots denoting individual cases extending further beyond the whiskers, provide a measure of outward dispersions and outlier cases—offenders with relatively large or extreme number of prior arrest records relative to the norm within each subgroup.

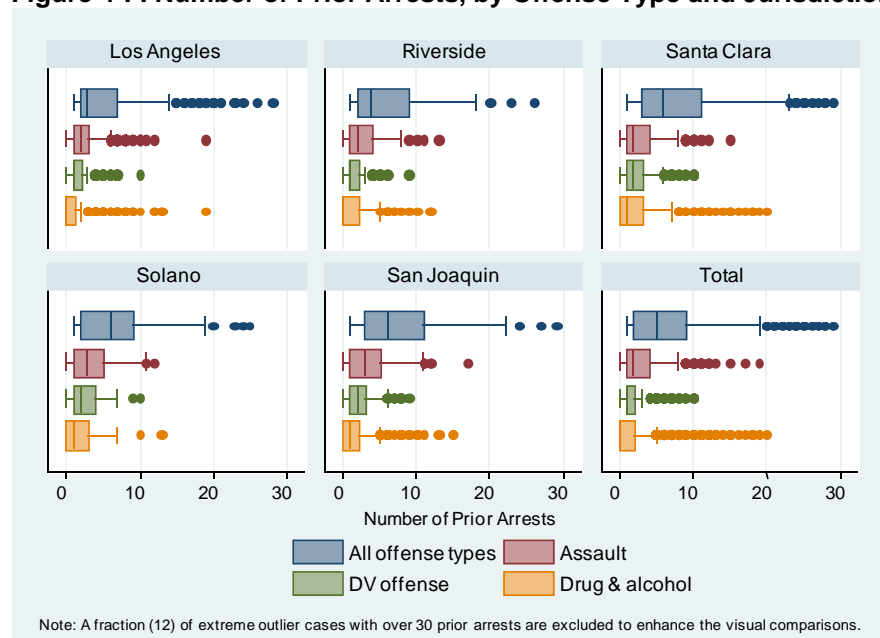
A few salient observations from Figure 4-E are summarized below:

- The two age groups under 30 are characterized by smaller dispersions (as represented by box width and whisker length) as well as lower median values, in clear distinction from their older counterparts in the two age groups over 30.
- While the dispersions in older age groups stretch out farther into the higher end in the continuum of prior arrest records—indicating an increasingly larger proportion of chronic offenders—there exists a sizable proportion of low-level offenders in all age groups, consistent with the discussion above regarding first-time offenders.
- In terms of greater dispersions of prior arrest frequencies being associated with older age groups, the association appears to be weaker in Los Angeles and Riverside compared with Santa Clara and San Joaquin. In other words, prior arrest frequencies show a stronger positive correlation with offender age in Santa Clara and San Joaquin than in Los Angeles and Riverside, suggesting the existence of some structurally persistent factors contributing to different offender profiles among the jurisdictions.

**Figure 4-E. Number of Prior Arrests of All Offense Types, by Age Group and Jurisdiction**



**Figure 4-F. Number of Prior Arrests, by Offense Type and Jurisdiction**



When the focus of analysis is shifted to specific offense types from total offenses, as depicted above, different patterns of offender profiles persist among jurisdictions. Figure 4-F shows the dispersion of prior arrests by jurisdiction, with offenses grouped into four categories: all offenses, assaults, domestic violence as a subset of overall assaults, and offenses related to drug and alcohol charges. Measured by any of these indicators, offenders in Los Angeles and Riverside consistently show fewer prior arrests relative to Santa Clara and San Joaquin. The different patterns remain unchanged even when controlling for potential confounding factors including race/ethnicity, current age, and non-English-speaking status of the offenders. The offender profile in Solano is relatively more difficult to ascertain because of its smaller sample size.

### Drug/Alcohol Abuse as Measured by CAGE Score

An offender's level of drug/alcohol abuse is closely correlated with arrest history. Higher scores on the drug/alcohol risk-assessment instrument correspond to a greater likelihood of previous arrest. The CAGE assessment, administered at intake, was used to screen for potential drug/alcohol abuse among study enrollees (see Table 4-C).

**Table 4-C. CAGE Score, by Jurisdiction**

| Score         | Los Angeles |     | Riverside  |     | Santa Clara |     | Solano    |     | San Joaquin |     | Total        |     |
|---------------|-------------|-----|------------|-----|-------------|-----|-----------|-----|-------------|-----|--------------|-----|
|               | N           | %   | N          | %   | N           | %   | N         | %   | N           | %   | N            | %   |
| 0             | 201         | 49% | 74         | 46% | 122         | 38% | 35        | 49% | 96          | 47% | 528          | 45% |
| 1             | 62          | 15% | 24         | 15% | 54          | 17% | 6         | 8%  | 29          | 14% | 175          | 15% |
| 2             | 71          | 17% | 24         | 15% | 60          | 19% | 13        | 18% | 36          | 17% | 204          | 18% |
| 3             | 43          | 10% | 18         | 11% | 48          | 15% | 9         | 13% | 26          | 13% | 144          | 12% |
| 4             | 33          | 8%  | 20         | 13% | 33          | 10% | 8         | 11% | 19          | 9%  | 113          | 10% |
| <b>Total</b>  | <b>410</b>  |     | <b>160</b> |     | <b>317</b>  |     | <b>71</b> |     | <b>206</b>  |     | <b>1,164</b> |     |
| Average Score | 1.13        |     | 1.29       |     | 1.42        |     | 1.28      |     | 1.24        |     | 1.26         |     |

Table 4-C shows that a little less than one half of the study population reported a score of zero on the CAGE assessment, ranging from 38 percent in Santa Clara County to 49 percent in Los Angeles and Solano Counties. A score of one on the CAGE suggests some drug/alcohol abuse while a score of two or greater is considered to indicate a drug/alcohol problem with about 90 percent accuracy. Thus, slightly more than one half of the study population showed at least some signs of possible drug/alcohol abuse while 40 percent of offenders in the study reported a score of two or higher.<sup>2</sup>

A related indicator of drug/alcohol abuse is whether the study subject is currently or has previously been enrolled in a drug or alcohol treatment program. As part of the intake process, offenders were asked whether they were currently enrolled in a drug/alcohol treatment program or had been enrolled in one previously. Although there is a correlation between higher CAGE scores and higher enrollment rates (see Table 4-D), the substance abuse program enrollment rate seems low for those who self-report most or all of the signs of drug alcohol abuse. For respondents with a CAGE score of 4, the enrollment rates in drug/alcohol treatment programs ranged from a low of 21 percent in San Joaquin to a high of 58 percent in Santa Clara.

**Table 4-D. Percent of Offenders Enrolled in Drug/Alcohol Treatment Programs Within Last 12 Months, by CAGE Score and Jurisdiction**

| CAGE Score | Los Angeles | Riverside | San Joaquin | Santa Clara | Solano | Average |
|------------|-------------|-----------|-------------|-------------|--------|---------|
| 0          | 3%          | 4%        | 4%          | 8%          | 11%    | 5%      |
| 1          | 11%         | 8%        | 14%         | 26%         | 33%    | 17%     |
| 2          | 15%         | 13%       | 8%          | 25%         | 15%    | 17%     |
| 3          | 26%         | 22%       | 38%         | 38%         | 44%    | 33%     |
| 4          | 27%         | 45%       | 21%         | 58%         | 25%    | 38%     |

The co-occurrence of drug/alcohol abuse and domestic violence is a complete topic in itself and is beyond the scope of this research.<sup>3</sup> Nevertheless, the self-reported data on the CAGE and the variance across jurisdictions in the proportion enrolled in drug/alcohol treatment programs suggest a number of issues. On one hand, it may indicate the need for enhanced screening of drug/alcohol issues as part of the probation or BIP intake process. It may also signal the difficulty of addressing all treatment needs at once; for some offenders, particularly those trying to maintain stable employment, attending and paying for both a weekly BIP program and drug/alcohol treatment sessions at the same time may be too difficult.

### ***Revised Conflict Tactics Scale 2 (CTS2)***

The CTS2 questionnaire (see Appendix C) seeks to measure the extent to which certain behavior and tactics have been used by a couple during a conflict.<sup>4</sup> The self-administered questionnaire consists of 39 questions,<sup>5</sup> aligned along five subscales, which attempt to assess a client's behavior in the following areas:

- Negotiation
- Psychological aggression
- Physical assault
- Injury

- Sexual coercion

The CTS2 was administered at program intake to 1,457 study enrollees. Of those, 1,237 surveys contained enough valid responses for analysis; 220 surveys had to be omitted because the respondent failed to answer enough questions pertaining to a subscale to generate a valid score.

Even for the 1,237 cases with “valid” responses that allowed for the calculation of subscale scores, there appears to be considerable suppression effects in the responses, leading to an overall underestimation of the prevalence of abusive behavior among the study subjects. There are likely to be multiple sources responsible for suppression effects in the responses, including social desirability (especially given the intrusive nature of some of the questions) and denial (from both psychological and legal perspectives). Furthermore, the self-administered survey may present considerable cognitive challenges for the respondents—who, as noted above, are generally limited in their educational attainment—that they are likely to exhibit “satisficing” behavior in completing the questionnaire; that is, putting in only minimally required effort in answering the questions. This would further affect the validity of the data.

An analysis of the survey data suggests the existence of suppression effects in the responses. The cross-item variance of responses for each of the subscales was calculated to help determine whether the responses were valid. To illustrate, a respondent who answered every question with “never” was probably not responding honestly, because some questions in a subscale are worded so that a response of “no” or “never” is the desired answer, and some are worded so that an answer of “no” or “never” is a very undesirable response. The variance across each of the subscales was calculated for each respondent. Some variance is expected as a result of the wording of the questions on the survey, so the surveys with little to no variance are probably invalid responses.

As another way to determine whether surveys were filled out truthfully, BIP staff were asked to flag intake materials, including the CTS2 survey, that appeared to be inaccurate. A comment box was provided for program staff to record their observations, such as “client circled zero on all questions” or “client reports he has never been violent in questionnaire but police report indicates [victim] has bruises on both of her eyes/cheeks.” Based on observations made by BIP staff administering the survey, some respondents might have been uncomfortable revealing information about past acts of violence because they mistakenly believed that the information might be used to prosecute them for another crime or could affect the status of their current case.

With these caveats on data limitations, the scores for each subscale were calculated and are shown in Table 4-E. When reviewing these scores it is important to remember that for the negotiation scale, higher scores indicate the reporting of more positive interactions with the partner. For the other four scales, higher scores are associated with more negative interactions.



**Table 4-E. CTS2 Scores by Subscale, by Jurisdiction**

| <b>County</b> | <b>Negotiation**</b> | <b>Psychological Aggression*</b> | <b>Physical Aggression</b> | <b>Injury (of Self) Inflicted by Partner</b> | <b>Sexual Coercion</b> |
|---------------|----------------------|----------------------------------|----------------------------|----------------------------------------------|------------------------|
| Los Angeles   | 57.47                | 22.22                            | 7.25                       | 2.68                                         | 3.32                   |
| Riverside     | 62.91                | 27.12                            | 7.70                       | 2.30                                         | 3.30                   |
| Santa Clara   | 66.05                | 24.84                            | 5.93                       | 2.19                                         | 1.86                   |
| Solano        | 72.14                | 30.58                            | 7.62                       | 3.05                                         | 2.73                   |
| San Joaquin   | 57.97                | 23.56                            | 7.00                       | 2.02                                         | 2.82                   |
| <b>Total</b>  | <b>61.53</b>         | <b>24.35</b>                     | <b>6.93</b>                | <b>2.39</b>                                  | <b>2.79</b>            |

\*Statistically significant differences at 5%; \*\*statistically significant at 1%

The data shows that while respondents were more likely to report acts of psychological aggression against their partners, they were less likely to self-report behavior that resulted in physical injury. This pattern may reflect less willingness on the part of clients to disclose severe forms of abuse or it may indicate that severe forms of abuse are less common among the client population—or both.

## CTS2 Prevalence Scores

In order to overcome the likelihood that clients underreport the frequency of abusive incidents, another way to examine the CTS2 data is to convert the five subscales into prevalence scores. Prevalence scores are useful primarily when certain behavior is less common or where there is good reason to believe it is underreported. They represent the proportion of a client population that reports *any* occurrence of abusive behavior as assessed within each CTS2 subscale. Prevalence scores of this kind tend to be much less prone to response bias, and they tend to be more accurate indicators of how common abuse is among the current client group.

Table 4-F displays the prevalence scores for four of the subscales on the CTS2 that assess some form of abuse. Each subscale is also broken into minor and severe DV incidents.

**Table 4-F. CTS2 Prevalence Scores, by Jurisdiction**

| <b>County</b> | <b>Psychological Aggression</b> |               | <b>Physical Aggression</b> |               | <b>Injury (of Self) Inflicted by Partner</b> |               | <b>Sexual Coercion</b> |               |
|---------------|---------------------------------|---------------|----------------------------|---------------|----------------------------------------------|---------------|------------------------|---------------|
|               | <i>Minor</i>                    | <i>Severe</i> | <i>Minor</i>               | <i>Severe</i> | <i>Minor</i>                                 | <i>Severe</i> | <i>Minor</i>           | <i>Severe</i> |
| Los Angeles   | 89%                             | 52%           | 71%                        | 39%           | 36%                                          | 13%           | 23%                    | 10%           |
| Riverside     | 88%                             | 59%           | 64%                        | 40%           | 41%                                          | 11%           | 20%                    | 5%            |
| Santa Clara   | 90%                             | 53%           | 68%                        | 38%           | 38%                                          | 11%           | 15%                    | 8%            |
| Solano        | 90%                             | 50%           | 66%                        | 34%           | 39%                                          | 10%           | 18%                    | 6%            |
| San Joaquin   | 87%                             | 57%           | 63%                        | 36%           | 33%                                          | 9%            | 15%                    | 9%            |
| <b>Total</b>  | <b>89%</b>                      | <b>54%</b>    | <b>67%</b>                 | <b>38%</b>    | <b>37%</b>                                   | <b>11%</b>    | <b>19%</b>             | <b>8%</b>     |

The data in Table 4-F suggests the following:

- Only about 10 percent of all respondents denied any abusive incidents at the intake session.
- Minor forms of abuse are more frequently reported than major forms, but it is striking how common some form of abuse is across clients, regardless of county of residence.
- The distribution of most forms of abuse does not vary by county. The exception to this rule can be seen in the sexual coercion scale. For this form of domestic abuse, there are statistically significant differences among county prevalence scores for minor forms of sexual coercion.

Despite these findings, overall the CTS2 data is of relatively limited value in predicting outcomes because of the high likelihood of suppressed responses and the extreme skewed distribution of the data.

## **Summary**

In this chapter we present data on over 1,000 domestic-violence offenders in our sample and examine the profiles of these men. These profiles will serve as controls to ensure the validity of the system-level analysis. We look at three major types of offender data: family relations, socio-economic status, criminal history and alcohol/drug abuse. Two supplemental instruments, the CAGE and CTS2, were also administered to study enrollees to further assess behavior and predict outcomes. Statistically-significant differences were found for certain characteristics across the three types of offender data, though the predictive power of the CTS2 data is low due to response suppression and the skewed distribution of the data. The importance of these offender characteristics becomes more apparent in Chapter 5 where we use the data on individual characteristics as part of the statistical models to evaluate two principal outcome measures: program completion and re-offense.

## Endnotes Chapter 4

---

1. L. W. Sherman and D. A. Smith, "Crime, Punishment, and Stake in Conformity: Legal and Informal Control of Domestic Violence" (1992) *57 American Sociological Review* 680–690.
2. See "Alcohol and Substance Abuse Evaluation," <http://emedicine.medscape.com/article/805084-overview>, Amy Cohagan, Richard Worthington (2007).
3. Lisa Lightman and Francine Byrne, "Addressing the Co-occurrence of Domestic Violence and Substance Abuse" (2005) *Journal of the Center of Families, Children and the Courts* 53-72.
4. Straus, Hamby, Boney-McCoy, and Sugarman (1996).
5. The original CTS and CTS2 questionnaires are designed to measure the behavior of the respondent and the respondent's partner (from the point of view of the respondent). For the purposes of this study, the survey instrument used in the present study included only the questions directed at the respondent.

## Chapter 5: Analysis of System Impacts

### ***Introduction***

For domestic violence offenders convicted and ordered to attend a 52-week BIP, successful program completion is one key outcome variable. Indeed, a common criticism of the justice system response to domestic violence and of evaluations of BIPs relates to the fact that such a small percentage of offenders actually complete the programs.<sup>1</sup> Program completion, however, is largely a means to an end. The ultimate goals of the criminal sanction, especially participation in a BIP, are to hold the offender accountable and to increase victim safety by preventing re-offense. To analyze the system impacts on domestic violence offenders across jurisdictions, therefore, we focus primarily on two outcome measures: (1) program completion and (2) re-offense. With a somewhat more limited data set we also look at the impact of the BIPs on offenders' attitudes and beliefs.

### ***Program Completion and Termination***

Data sources for program completion came from participating BIPs providing updates on their clients' status on a regular basis. One immediate question that presented itself in evaluating the data had to do with the meaning of "program termination." At first glance it appeared that program termination could be treated simply as the opposite of successful program completion. The dichotomy between termination and completion, however, misrepresents the reality in which it is not uncommon to see termination followed by reenrollment—with the possibility of the sequence repeated more than once for some individuals—and subsequently the possibility of successful completion.

Variations across counties and across BIPs in policies and practices with regard to termination, re-enrollment, and whether prior attendances are given credit should all play an important role in determining the offender's ultimate chance of completing the program. Given the varying patterns across the jurisdictions regarding the sequence of termination and reenrollment, it becomes necessary to examine program completion rates in conjunction with termination patterns.

Table 5-A shows the number and percentage of offenders who either completed the 52-week program by the end of data collection in early February 2008 or had a record of having been terminated regardless of whether they were subsequently reenrolled in the program following the termination. The first thing to note in the table is that there are more cases with available data for tracking and analyzing terminations than program completions. This is largely due to cases that had an earlier termination record, but updated information indicated that the offenders were still active in the program as of the end of data collection.<sup>2</sup>

In general, completion and termination rates across the jurisdictions present a fairly consistent picture: jurisdictions with higher completion rates relative to another jurisdiction tend to show lower termination rates as well. The exceptions are the results in Santa Clara and San Joaquin.

While San Joaquin shows the highest termination rate at 64 percent compared with 55 percent in Santa Clara, there is no significant difference in their completion rates—52 percent in San Joaquin and 50 percent in Santa Clara.

**Table 5-A. Program Completion and Termination at End of Data Collection**

| Jurisdiction | Successful Completion |                       |                 | Termination  |                        |                  |
|--------------|-----------------------|-----------------------|-----------------|--------------|------------------------|------------------|
|              | Total Sample          | Number of Completions | Completion Rate | Total Sample | Number of Terminations | Termination Rate |
| Los Angeles  | 422                   | 237                   | 56%             | 471          | 209                    | 44%              |
| Riverside    | 145                   | 90                    | 62%             | 158          | 65                     | 41%              |
| Santa Clara  | 376                   | 187                   | 50%             | 399          | 220                    | 55%              |
| Solano       | 86                    | 56                    | 65%             | 86           | 35                     | 41%              |
| San Joaquin  | 227                   | 117                   | 52%             | 272          | 173                    | 64%              |
| Total        | 1,256                 | 687                   | 55%             | 1,386        | 702                    | 51%              |

Note: Without controlling for confounding factors, differences across jurisdictions are statistically significant at 1 percent level for both outcome measures.

In addition to the termination rates shown in Table 5-A as a static, end-point measure, Figure 5-A displays a dynamic picture of the growing proportion of program clients terminated over a span of 400 days from program intake. With the passage of time (represented on the horizontal axis), each client who is terminated from a program at a different time after enrollment adds to the total number of terminations, contributing to the upward movement of the lines depicting the different rates of termination in each jurisdiction.

**Figure 5-A. Cumulative Program Termination Rate from Intake to 400 Days**

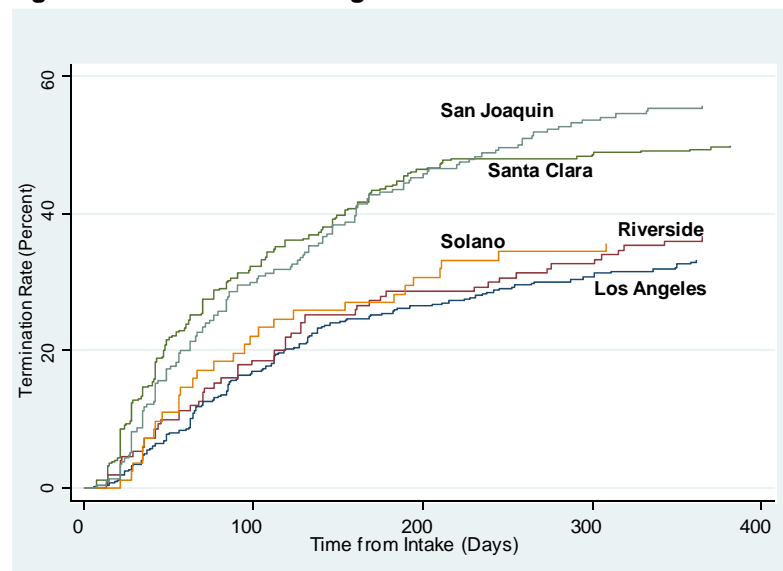


Figure 5-A helps to show how patterns of termination change over time. Although the *final* program outcomes in Santa Clara and San Joaquin vary—as Table 5-A reveals—Figure 5-A shows that the two jurisdictions actually follow a very similar trajectory over a substantial

duration of the follow-up period. It is only approximately 250 days after program enrollment that the patterns of program termination begin to diverge. Survival analysis of the trend lines shows no statistically significant difference between Santa Clara and San Joaquin. The other three jurisdictions (Los Angeles, Solano, and Riverside) show similar paths throughout the entire follow-up period, in clear distinction from the patterns in Santa Clara and San Joaquin.

## **Descriptive Analysis of Risk Factors**

Before delving into the analysis of system impacts or teasing out the relationship between termination and completion, this section provides a descriptive analysis of major risk factors that were found to have a strong correlation with program terminations and completions. The analysis is largely descriptive in nature, as the potential confounding effects of each of the risk factors in interaction with one another are not controlled for. For example, if the data reveals a strong correlation between offender-victim relationship and program completion, this connection might result from the effect of offender age as an intervening factor because offender age is correlated with both offender-victim relationship and program completion. When the potential confounding effect of offender age is controlled for, what was attributed to the positive impact of marital status might prove to be either exaggerated or merely the result of a spurious correlation with offender age.

Similar confounding effects are likely to exist in various other socioeconomic and demographic variables examined in this section. Following this descriptive analysis, the next section discusses results from various regression models in which the issue of confounding effects is specifically addressed. Regression analysis is also the statistical technique employed to address the central questions of the study, including the analysis of system-level impacts and the role of different justice system interventions in the overall system-level impacts.

Risk factors examined in Table 5-B are organized into the following categories, similar to those presented previously in the offender profile section:

- Family relations, including offender relationship with the victim and children;
- Current (as of program enrollment) employment and income level;
- Socio-demographic factors; and
- Criminal history and drug/alcohol abuse.

## **Relationship with Victim and Children**

Without considering possible interactions with confounding factors, Table 5-B shows that an offender has a noticeably higher chance of successfully completing the program if he is living with the victim at the time of program enrollment, if he has children and is living with them, and if the victim is his wife—either current or former—as opposed to a girlfriend. With the difference in program completion between the various comparison groups in the range of 10 to 15 percentage points, the impacts of these family-relations variables appear to be important risk/protective factors, as discussed in Chapter 4, on offender profiles.

**Table 5-B. Risk Factors and Program Completion Rate**

|                                           | Number of<br>Cases | Completion<br>Rate |                                        | Number of<br>Cases | Completion<br>Rate |
|-------------------------------------------|--------------------|--------------------|----------------------------------------|--------------------|--------------------|
| <u>Lives with Victim</u>                  |                    |                    | <u>Age at Intake</u>                   |                    |                    |
| No                                        | 742                | 51%                | <25                                    | 254                | 49%                |
| Yes                                       | 467                | 61%                | 25 - 29                                | 231                | 50%                |
| Total                                     | 1,209              | 55%                | 30 - 39                                | 358                | 56%                |
|                                           |                    |                    | >= 40                                  | 304                | 63%                |
| <u>Victim Is Wife-- Current or Former</u> |                    |                    | Total                                  | 1,147              | 55%                |
| No                                        | 674                | 48%                | <u>CAGE Score (drug/alcohol abuse)</u> |                    |                    |
| Yes                                       | 540                | 63%                | 0                                      | 438                | 57%                |
| Total                                     | 1,214              | 55%                | 1                                      | 147                | 54%                |
| <u>Lives with Children</u>                |                    |                    | 2                                      | 174                | 48%                |
| No children                               | 316                | 51%                | 3                                      | 128                | 38%                |
| Lives with children                       | 402                | 63%                | 4                                      | 99                 | 45%                |
| Visits children regularly                 | 301                | 49%                | Total                                  | 986                | 52%                |
| Does not visit children regularly         | 170                | 50%                | <u>Age at First Arrest</u>             |                    |                    |
| Total                                     | 1,189              | 54%                | <18                                    | 186                | 40%                |
| <u>Lost Job in Past Year</u>              |                    |                    | 18 - 24                                | 543                | 52%                |
| No                                        | 773                | 53%                | 25 - 29                                | 147                | 63%                |
| Yes                                       | 216                | 44%                | 30 - 39                                | 160                | 63%                |
| Total                                     | 989                | 51%                | >=40                                   | 97                 | 81%                |
| <u>Employment Status</u>                  |                    |                    | Total                                  | 1,133              | 55%                |
| Employed full-time                        | 472                | 60%                | <u>Total Prior Arrests</u>             |                    |                    |
| Employed part-time                        | 162                | 51%                | 1                                      | 210                | 76%                |
| Not employed                              | 367                | 41%                | 2 - 3                                  | 258                | 65%                |
| Total                                     | 1,001              | 51%                | 4 - 5                                  | 161                | 55%                |
| <u>Income</u>                             |                    |                    | >=6                                    | 506                | 42%                |
| \$0 - \$4,999                             | 284                | 48%                | Total                                  | 1,135              | 55%                |
| \$5,000 - \$14,999                        | 264                | 47%                | <u>Prior Assault Arrests</u>           |                    |                    |
| \$15,000 - \$24,999                       | 260                | 59%                | 1                                      | 443                | 67%                |
| \$25,000 - \$39,999                       | 211                | 61%                | 2                                      | 221                | 57%                |
| >=\$40,000                                | 106                | 68%                | 3-4                                    | 233                | 49%                |
| Total                                     | 1,125              | 54%                | >=5                                    | 238                | 38%                |
| <u>Education: Some College or More</u>    |                    |                    | Total                                  | 1,135              | 55%                |
| No                                        | 911                | 52%                | <u>Prior DV Arrests</u>                |                    |                    |
| Yes                                       | 271                | 62%                | 1                                      | 608                | 62%                |
| Total                                     | 1,182              | 55%                | 2                                      | 252                | 53%                |
| <u>Race/Ethnicity</u>                     |                    |                    | >=3                                    | 275                | 43%                |
| African American                          | 193                | 39%                | Total                                  | 1,135              | 55%                |
| Hispanic                                  | 615                | 60%                | <u>Prior Drug Arrests</u>              |                    |                    |
| White                                     | 245                | 50%                | No                                     | 559                | 64%                |
| Other                                     | 127                | 69%                | Yes                                    | 576                | 47%                |
| Total                                     | 1,180              | 55%                | Total                                  | 1,135              | 55%                |
| <u>Non-English Speaker</u>                |                    |                    | <u>Prior Felony Arrests</u>            |                    |                    |
| No                                        | 963                | 50%                | No                                     | 122                | 66%                |
| Yes                                       | 293                | 69%                | Yes                                    | 1,013              | 54%                |
| Total                                     | 1,256              | 55%                | Total                                  | 1,135              | 55%                |

Note: Differences across comparison groups in completion rates are all statistically significant at a minimum of 5% level based on ANOVA.

## Employment and Income

Consistent with the results associated with family-relations factors above, the financial condition of the offenders as measured by employment status and income appears to point to the positive

impact of a more stable living situation on program completion. Offenders who reported that they were fully employed at the time of program enrollment—about 50 percent of the total sample—had a completion rate of 60 percent. Compared with this group, the program completion rate for those employed only part-time is lower by 9 percentage points. The rate drops by an additional 10 percentage points to 41 percent for offenders who are unemployed, which accounts for approximately 35 percent of the total sample.

A slightly different measure of economic stability—job loss in the past year—shows similar impacts on program completion. Those who experienced recent job loss had a completion rate of approximately 45 percent, again about 10 percent lower than those who did not experience it.

The negative impact of economic disadvantage on program completion is also evident as measured by self-reported income by the offenders. With annual income grouped into five categories, the data suggests a nonlinear income effect, with two relevant threshold points affecting program completion rates. Offenders whose self-reported income fell into the first and second intervals—below \$5,000 per year, and between \$5,000 and \$14,999—had very similar completion rates, approximately 50 percent. As the income of offenders increased in the third and fourth intervals—between \$15,000 and \$24,999 and between \$25,000 and \$39,999—the corresponding completion rate rises by about 10 percentage points, to approximately 60 percent. Those in the highest income group in the sample—above \$40,000, representing about 10 percent of the total sample—had a completion rate of almost 70 percent.

In addition to self-reported income, another proxy measure of income, the amount of fees charged for program attendance, reveals similar impacts on program completion. For those charged less than \$20 per week (about 15 percent of the total sample), approximately 40 percent successfully completed the program, compared with a completion rate of 50 percent for those charged between \$20 to \$25 per week (representing 20 percent of the total sample). For the more than 60 percent of the program clients charged over \$25 per week, the completion rate rises to nearly 60 percent.

The strength of the connection between fees and completion rate varies to some extent among the jurisdictions. The clear exception is Santa Clara County, where the correlation is almost nonexistent at fee levels under \$40. However, the approximately 15 percent of clients in Santa Clara charged over \$40 had a completion rate of 75 percent compared with a rate of about 45 percent for the rest of the clients.

### **Socio-demographic Factors**

Socio-demographic variables that revealed significant correlation with completion rate include education, race/ethnicity, whether the offender is identified as a non-English speaker, and age at intake.

Years of educational attainment, aggregated into two groups (offenders who completed only high school or less and those who attended some college or more), have a strong positive correlation with program completion. There is a 10 percent difference in completion rates between the two groups: 52 percent for high school graduates versus 62 percent for those with some college or more.



A comparison by race/ethnicity shows differences in completion rates ranging from a high of 69 percent for the “other” group and a low of 39 percent for African Americans. Both of these groups account for a minority proportion of the total sample: 11 percent for “others” and 16 percent for African Americans. As the largest group in the sample (slightly more than half of the total), Hispanics completed the programs at a rate of 60 percent. Whites, who make up 21 percent of the total sample, showed a completion rate of 50 percent.

As noted in Chapter 4, on offender profiles, approximately 35 percent of Hispanics and “others” are non-English-speaking, and they appear to exhibit different profiles compared with their English-speaking counterparts. Analysis results of completion rates provide noticeably higher completion rates for non-English speakers. For Hispanics, the difference in completion rates between English and non-English speakers is 12 percentage points: 55 percent for the former and 67 percent for the latter. The difference between the two subgroups for those in the “other” race/ethnicity category is more than 15 percentage points, with 82 percent of non-English speakers completing the program compared with 65 percent of English-speaking offenders.

For the four age-at-intake groups presented in Table 5-B, there is virtually no difference in completion rate between those under the age of 25 and those between 25 and 29; both completed the program at a rate of about 50 percent. The completion rate rises slightly to 56 percent for those between the ages of 30 and 39. For those above 40—representing about one-quarter of the total sample—the completion rate rises further to 63 percent.

### **Criminal History and Drug/Alcohol Abuse**

All measures of criminal history in Table 5-B show a consistently negative impact on completion rate. In general, those with earlier onset of criminal activities and recording a higher number of prior arrests (regardless of offense type) are less likely to complete the program.

Among offenders whose first arrest occurred before the age of 18, only 40 percent completed the program successfully. The completion rate rises to 52 percent for those first arrested between the ages 18 and 24; it rises further to 63 percent for the next two age groups (25 to 29 and 30 to 39). For those who experienced their first arrest over the age of 40, 81 percent completed the program successfully.

With prior arrest records grouped into various offense types, the five different measures for prior arrests in Table 5-B show similar patterns of correlation with completion rate. Without differentiating offense types, data shows that first offenders—those with one prior arrest—completed the program at a rate of 76 percent. In contrast, the completion rate is only 42 percent for those with more than six prior arrests. In between the two groups, it appears two additional priors are associated with a decline of approximately 10 percentage points in completion rate: declining from 76 percent for first offenders to 65 percent for those with 2 to 3 prior arrests, and dropping further to 55 percent for those with 4 to 5 prior arrests. Other prior arrest measures limited to specific offense types—including offenses involving assault, domestic violence, drug, or felony charges—all show similarly strong correlations with program completion rate.

Similar to the predictive power of prior drug arrests in relation to program completion, it is important to note that CAGE scores, as a simple but highly effective tool for initial assessment of drug and alcohol problems, also indicate a strong correlation between drug/alcohol problems and program completion. For approximately 45 percent of the offenders whose CAGE score revealed no clear sign of drug/alcohol issues (score=0), nearly 57 percent completed the program. As the CAGE score rises, suggesting more severe problems with drug and alcohol, the completion rate tends to decline.

## **Regression Analysis of Program Completion and Termination**

Given the correlations between the variables described above and program completion, these risk factors serve as control variables in regression analyses in the following section. To the extent these control variables represent the relevant factors that affect the offender's propensity for program completion, regression models provide a means of answering the central questions of this study.

- Do system impacts vary significantly across jurisdictions?
- Do the impacts vary systematically across programs within jurisdictions?
- To what extent do variances at the program level account for differences in jurisdictional effects?

It is important to note that the term “effect” used in the following discussion should be considered within the context of the study design, specifically, the scheme of comparison in examining different outcome variables. There are comparisons across jurisdictions; there are also comparisons regarding the relative impact from the program versus probation and court supervision at the jurisdictional level. The common element in all comparisons consists of batterers enrolled in 52-week programs and under some form of probation or court supervision. When referring to “program effect,” therefore, the analysis does not attempt to answer the question of whether the programs are working in holding the batterers accountable compared with other intervention strategies in which the offenders are not enrolled in 52-week programs.

## **Regression Method**

Table 5-C shows regression results for *program completion* based on two basic models. These models seek to disentangle the effects of multiple variables operating at different levels—individual offenders with different profiles, BIPs that vary in their policies and treatment models, and the jurisdictional effect of court and probation oversight. Models 1 to 4 use regular regression for handling dependent variables with binary outcomes (yes or no for program completion). In these basic models, program-level effects are treated as an estimation problem related to the fundamental assumptions of regression analysis.<sup>3</sup> Without appropriately addressing the interactions between individual offenders and programs in their nested structure, regular regression models may lead to biased or misleading results concerning system-level impacts.

As a base comparison model, model 1 in Table 5-C includes jurisdictions as the only independent variables. This model does not control for any confounding factors at the level of

individual offenders or correct any estimation problems noted above. Models 2 to 4 enhance the base model successively in the following ways by

- improving the accuracy of significance testing (model 2);
- adding a subset of demographic factors to the model to test the stability of the model in response to the inclusion of control variables as well as changes in sample size caused by missing data in control variables (model 3), and;
- presenting a full model with all relevant control variables available included in the model (model 4).<sup>4</sup>

Models 5 to 7 are specified in the same manner as models 2 to 4 with regard to the set of control variables included, but variances at the program level are examined explicitly using a special multilevel regression model.<sup>5</sup> In addition to estimates of jurisdictional effects (as represented by a series of binary variables), multilevel models also provide an estimate of program-level variances, along with standard errors for testing the statistical significance of the estimate.

### **Findings from Regression Analysis**

Comparison of the various regression models shows there are significant confounding effects that arise from both program- and individual offender-level variables. This can be seen in the steady decline of statistical significance associated with the jurisdiction variables as additional variables are added to the base model in the regular models (2 to 4). The same declining significance can also be seen in the estimates of program-level effects in the multilevel models (5 to 7).

What appear at first glance in Table 5-A to be statistically significant differences in completion rates across the jurisdictions can be attributed largely to some systematic variances across the programs and differences in offender profiles. In other words, these regression models provide no strong evidence that the different intervention strategies at the system level have any impact on BIP completion rates independent of program-level effects and offender profile differences. The only exception is in Solano County, where results suggesting a higher completion rate cannot be explained away by the observable confounding factors included in the models.

**Table 5-C. Regression Results for Program Completion**

| Dependent Variable:<br>Program Completion                                                   | Logistic Regression Models |                 |                   |                   | Multilevel Logistic Regression Models |                   |                   |
|---------------------------------------------------------------------------------------------|----------------------------|-----------------|-------------------|-------------------|---------------------------------------|-------------------|-------------------|
|                                                                                             | (1) <sup>†</sup>           | (2)             | (3)               | (4)               | (5)                                   | (6)               | (7)               |
| <i>Jurisdiction (Los Angeles as the base comparison group)</i>                              |                            |                 |                   |                   |                                       |                   |                   |
| Riverside                                                                                   | 1.277<br>(1.24)            | 1.277<br>(0.79) | 1.511<br>(1.40)   | 1.498<br>(1.64)   | 1.155<br>(0.32)                       | 1.404<br>(0.82)   | 1.456<br>(1.08)   |
| Santa Clara                                                                                 | 0.772<br>(1.81)            | 0.772<br>(1.01) | 0.840<br>(0.66)   | 0.886<br>(0.47)   | 0.991<br>(0.02)                       | 0.981<br>(0.06)   | 0.958<br>(0.15)   |
| Solano                                                                                      | 1.457<br>(1.53)            | 1.457<br>(1.52) | 2.007<br>(1.87)   | 3.709<br>(2.61)** | 1.896<br>(1.17)                       | 2.683<br>(1.91)   | 3.895<br>(2.54)*  |
| San Joaquin                                                                                 | 0.830<br>(1.13)            | 0.830<br>(0.61) | 1.092<br>(0.30)   | 1.233<br>(0.87)   | 1.076<br>(0.17)                       | 1.315<br>(0.71)   | 1.265<br>(0.71)   |
| <i>No children as base comparison group</i>                                                 |                            |                 |                   |                   |                                       |                   |                   |
| Lives with children                                                                         |                            |                 | 1.286<br>(1.54)   | 1.088<br>(0.40)   |                                       | 1.331<br>(1.62)   | 1.115<br>(0.50)   |
| Visits children regularly                                                                   |                            |                 | 0.865<br>(0.98)   | 0.804<br>(0.86)   |                                       | 0.957<br>(0.24)   | 0.829<br>(0.85)   |
| Does not visit children regularly                                                           |                            |                 | 0.884<br>(0.43)   | 0.810<br>(0.59)   |                                       | 0.909<br>(0.44)   | 0.806<br>(0.79)   |
| Education: some college or more                                                             |                            |                 | 1.607<br>(2.25)*  | 1.543<br>(1.65)   |                                       | 1.693<br>(3.28)** | 1.574<br>(2.20)*  |
| Victim is wife--former and current                                                          |                            |                 | 1.551<br>(3.53)** | 1.398<br>(1.76)   |                                       | 1.456<br>(2.69)** | 1.353<br>(1.69)   |
| Non-English speaker                                                                         |                            |                 | 2.247<br>(4.70)** | 1.061<br>(0.32)   |                                       | 2.143<br>(4.23)** | 1.089<br>(0.35)   |
| <i>Employment Status (employed full-time as base comparison group)</i>                      |                            |                 |                   |                   |                                       |                   |                   |
| Employed part-time                                                                          |                            |                 |                   | 1.007<br>(0.03)   |                                       |                   | 0.975<br>(0.11)   |
| Not employed                                                                                |                            |                 |                   | 0.785<br>(1.43)   |                                       |                   | 0.780<br>(1.33)   |
| <i>Race/ethnicity (African American as base comparison group)</i>                           |                            |                 |                   |                   |                                       |                   |                   |
| Hispanic                                                                                    |                            |                 |                   | 2.613<br>(2.93)** |                                       |                   | 2.717<br>(3.96)** |
| Other                                                                                       |                            |                 |                   | 1.926<br>(1.46)   |                                       |                   | 2.014<br>(1.93)   |
| White                                                                                       |                            |                 |                   | 1.629<br>(1.27)   |                                       |                   | 1.664<br>(1.88)   |
| <i>Total prior arrests for any offense (single, first offense as base comparison group)</i> |                            |                 |                   |                   |                                       |                   |                   |
| 2-3                                                                                         |                            |                 |                   | 0.797<br>(0.71)   |                                       |                   | 0.768<br>(0.87)   |
| 4-5                                                                                         |                            |                 |                   | 0.621<br>(1.44)   |                                       |                   | 0.601<br>(1.41)   |
| >=6                                                                                         |                            |                 |                   | 0.551<br>(1.70)   |                                       |                   | 0.540<br>(1.60)   |
| <i>Total prior arrests for DV offenses (single, first offense as base comparison group)</i> |                            |                 |                   |                   |                                       |                   |                   |
| 2                                                                                           |                            |                 |                   | 1.013<br>(0.07)   |                                       |                   | 1.002<br>(0.01)   |
| >=3                                                                                         |                            |                 |                   | 0.760<br>(0.95)   |                                       |                   | 0.751<br>(1.28)   |
| Had prior felony arrests                                                                    |                            |                 |                   | 0.782<br>(0.83)   |                                       |                   | 0.793<br>(0.78)   |
| Had prior arrests for drug offenses                                                         |                            |                 |                   | 0.700<br>(2.58)** |                                       |                   | 0.704<br>(1.79)   |
| Age at intake                                                                               |                            |                 |                   | 1.036<br>(3.32)** |                                       |                   | 1.036<br>(2.82)** |
| Age at first arrest                                                                         |                            |                 |                   | 0.999<br>(0.10)   |                                       |                   | 0.998<br>(0.10)   |
| <i>CAGE Score (0 as base comparison group)</i>                                              |                            |                 |                   |                   |                                       |                   |                   |
| CAGE = 1                                                                                    |                            |                 |                   | 0.953<br>(0.16)   |                                       |                   | 0.932<br>(0.30)   |
| CAGE = 2                                                                                    |                            |                 |                   | 0.668<br>(1.88)   |                                       |                   | 0.665<br>(1.78)   |
| CAGE = 3                                                                                    |                            |                 |                   | 0.444<br>(3.54)** |                                       |                   | 0.427<br>(3.27)** |
| CAGE = 4                                                                                    |                            |                 |                   | 0.705<br>(1.17)   |                                       |                   | 0.700<br>(1.29)   |
| Program level variance                                                                      |                            |                 |                   |                   | 0.619<br>(2.64)**                     | 0.424<br>(2.22)*  | 0.131<br>(1.01)   |
| Observations                                                                                | 1,256                      | 1,256           | 1,143             | 802               | 1,256                                 | 1,143             | 802               |

Absolute value of z statistics in parentheses; \* significant at 5%; \*\* significant at 1%.

<sup>†</sup> In models 2 to 4, Huber-White sandwich estimate of variance is used for robust estimates of standard errors that adjust for intraclass correlations at the level of batterer intervention programs.

**Table 5-D. Regression Results for Program Termination**

| Dependent Variable:<br>Program Termination                                                 | Logistic Regression Models |                   |                   |                   | Multilevel Logistic Regression Models |                   |                   |
|--------------------------------------------------------------------------------------------|----------------------------|-------------------|-------------------|-------------------|---------------------------------------|-------------------|-------------------|
|                                                                                            | (1) <sup>†</sup>           | (2)               | (3)               | (4)               | (5)                                   | (6)               | (7)               |
| <i>Jurisdiction (Los Angeles as the base comparison group)</i>                             |                            |                   |                   |                   |                                       |                   |                   |
| Riverside                                                                                  | 0.876<br>(0.71)            | 0.876<br>(0.42)   | 0.728<br>(1.06)   | 0.763<br>(1.09)   | 0.932<br>(0.17)                       | 0.747<br>(0.72)   | 0.778<br>(0.76)   |
| Santa Clara                                                                                | 1.541<br>(3.16)**          | 1.541<br>(1.89)   | 1.468<br>(1.79)   | 1.344<br>(1.44)   | 1.183<br>(0.48)                       | 1.187<br>(0.52)   | 1.222<br>(0.75)   |
| Solano                                                                                     | 0.860<br>(0.63)            | 0.860<br>(0.55)   | 0.526<br>(1.37)   | 0.357<br>(1.75)   | 0.641<br>(0.87)                       | 0.358<br>(2.00)*  | 0.339<br>(2.14)*  |
| San Joaquin                                                                                | 2.191<br>(5.01)**          | 2.191<br>(2.63)** | 1.726<br>(1.95)   | 1.753<br>(1.88)   | 1.902<br>(1.63)                       | 1.623<br>(1.30)   | 1.964<br>(2.08)*  |
| <i>No children as base comparison group</i>                                                |                            |                   |                   |                   |                                       |                   |                   |
| Lives with children                                                                        |                            |                   | 0.666<br>(2.60)** | 0.800<br>(1.08)   |                                       | 0.647<br>(2.57)*  | 0.781<br>(1.18)   |
| Visits children regularly                                                                  |                            |                   | 0.987<br>(0.12)   | 1.043<br>(0.21)   |                                       | 0.889<br>(0.68)   | 1.004<br>(0.02)   |
| Does not visit children regularly                                                          |                            |                   | 0.981<br>(0.07)   | 1.068<br>(0.22)   |                                       | 0.975<br>(0.12)   | 1.079<br>(0.29)   |
| Education: some college or more                                                            |                            |                   | 0.604<br>(2.66)** | 0.604<br>(2.27)*  |                                       | 0.560<br>(3.83)** | 0.577<br>(2.88)** |
| Victim is wife--former and current                                                         |                            |                   | 0.673<br>(3.61)** | 0.756<br>(1.63)   |                                       | 0.711<br>(2.58)** | 0.772<br>(1.53)   |
| Non-English Speaker                                                                        |                            |                   | 0.426<br>(5.63)** | 0.812<br>(1.24)   |                                       | 0.455<br>(4.60)** | 0.797<br>(0.99)   |
| <i>Employment Status (employed full-time as base comparison group)</i>                     |                            |                   |                   |                   |                                       |                   |                   |
| Employed part-time                                                                         |                            |                   |                   | 1.094<br>(0.46)   |                                       |                   | 1.109<br>(0.47)   |
| Not employed                                                                               |                            |                   |                   | 1.468<br>(2.23)*  |                                       |                   | 1.451<br>(2.10)*  |
| <i>Race/ethnicity (African American as base comparison group)</i>                          |                            |                   |                   |                   |                                       |                   |                   |
| Hispanic                                                                                   |                            |                   |                   | 0.464<br>(2.41)*  |                                       |                   | 0.441<br>(3.37)** |
| Other                                                                                      |                            |                   |                   | 0.544<br>(1.43)   |                                       |                   | 0.520<br>(1.90)   |
| White                                                                                      |                            |                   |                   | 0.658<br>(1.22)   |                                       |                   | 0.625<br>(1.79)   |
| <i>Total prior arrests of any offense (single, first offense as base comparison group)</i> |                            |                   |                   |                   |                                       |                   |                   |
| 2-3                                                                                        |                            |                   |                   | 1.510<br>(1.57)   |                                       |                   | 1.538<br>(1.55)   |
| 4-5                                                                                        |                            |                   |                   | 2.052<br>(2.00)*  |                                       |                   | 2.023<br>(2.10)*  |
| >=6                                                                                        |                            |                   |                   | 2.089<br>(2.13)*  |                                       |                   | 2.063<br>(1.98)*  |
| <i>Total prior arrests of DV offenses (single, first offense as base comparison group)</i> |                            |                   |                   |                   |                                       |                   |                   |
| 2                                                                                          |                            |                   |                   | 0.960<br>(0.19)   |                                       |                   | 0.963<br>(0.18)   |
| >=3                                                                                        |                            |                   |                   | 1.267<br>(0.95)   |                                       |                   | 1.298<br>(1.20)   |
| Had prior felony arrests                                                                   |                            |                   |                   | 1.068<br>(0.26)   |                                       |                   | 1.091<br>(0.31)   |
| Had prior arrests for drug offenses                                                        |                            |                   |                   | 1.238<br>(1.73)   |                                       |                   | 1.243<br>(1.15)   |
| Age at intake                                                                              |                            |                   |                   | 0.972<br>(2.52)*  |                                       |                   | 0.971<br>(2.42)*  |
| Age at first arrest                                                                        |                            |                   |                   | 1.001<br>(0.04)   |                                       |                   | 1.001<br>(0.09)   |
| <i>CAGE Score (0 as base comparison group)</i>                                             |                            |                   |                   |                   |                                       |                   |                   |
| CAGE = 1                                                                                   |                            |                   |                   | 0.990<br>(0.04)   |                                       |                   | 1.023<br>(0.10)   |
| CAGE = 2                                                                                   |                            |                   |                   | 1.532<br>(1.71)   |                                       |                   | 1.534<br>(1.96)   |
| CAGE = 3                                                                                   |                            |                   |                   | 1.820<br>(2.69)** |                                       |                   | 1.891<br>(2.56)*  |
| CAGE = 4                                                                                   |                            |                   |                   | 1.320<br>(0.96)   |                                       |                   | 1.354<br>(1.14)   |
| Program level variance                                                                     |                            |                   |                   |                   | 0.514<br>(2.76)**                     | 0.404<br>(2.42)*  | 0.128<br>(1.16)   |
| Observations                                                                               | 1,386                      | 1,386             | 1,270             | 892               | 1,386                                 | 1,270             | 892               |

Absolute value of z statistics in parentheses; \* significant at 5%; \*\* significant at 1%.

<sup>†</sup> In models 2 to 4, Huber-White sandwich estimate of variance is used for robust estimates of standard errors that adjust for intraclass correlations at the level of batterer intervention programs.

Relying on the same set of regression models but looking now at the impact of these variables on *program termination*, Table 5-D shows a somewhat different dynamic among the multiple levels of independent variables. In particular, results from both the regular and multilevel models indicate a statistically significant higher probability of termination in San Joaquin—in contrast to the lack of any significant difference in offenders’ completion rate in San Joaquin compared with other jurisdictions. In addition, the multilevel model shows that offenders in Solano are significantly less likely to be terminated from a BIP, consistent with the analysis of program completion in the previous model.

As described above, termination as an outcome measure is recorded as any incident of termination during the study period, regardless of the possibility in some cases of subsequent reenrollment and even eventual program completion. Differences across the jurisdictions in the prevalence of these cases may reflect different policies on sanctions against noncompliance.

In jurisdictions with more restrictive policies, a termination record may entail a more severe sanction such as probation revocation or incarceration, making it less likely for a case to reappear in the form of reenrollment and ultimately program completion. In jurisdictions where greater emphasis is placed on keeping offenders in a BIP, termination is less likely to translate into failure to complete the program. These different policy responses to program termination—extending to policies on absence, reenrollment, and credits for prior attendances—are essential to the system-level intervention that may produce the results described above for San Joaquin.

In addition to the different patterns of jurisdictional effects revealed by measures of completion versus termination, offender-level variables included in the regression models also show different patterns of association with the two outcome measures. Variables exhibiting stronger effects on termination, with weaker or no association with program completion, include living arrangements with children, employment status, prior arrest history, and prior arrest for drug charges. More specifically, an offender is less likely to experience termination if he has children and is living with them, is employed full-time, and has a less extensive prior arrest history. Offenders with prior drug arrests, in contrast, have a much lower probability of program completion, though only a slightly greater likelihood of program termination.<sup>6</sup>

Variables that show fairly stable and consistent impacts on both program termination and completion include relationship with victim, education, race/ethnicity, age at intake, and CAGE indicators of drug/alcohol abuse. Hispanics show a substantially higher chance of program completion compared with all other groups; older offenders have a greater chance of completing the program; higher CAGE scores—in particular, over the threshold value of 2—are associated with a lower probability of program completion; and finally, an offender whose victim is his wife (current or former), has a higher chance of completing the program than an offender whose victim is identified as a girlfriend.<sup>7</sup>

Some variables do not reveal any statistically significant impact on either completion or termination, including status as a non-English speaker, prior history of domestic violence arrests, and age at first arrest. The lack of statistical significance for these variables can be partly explained by the presence in the models of other correlated variables with more stable and robust

results. For example, the correlation between prior total arrests and domestic violence arrests is .47, and between the age at intake and age at first arrest .64.

Based on the different modes of analysis above—including descriptive analysis to examine the impact of individual risk factors, two types of regression models to address the nested structure of the multiple levels of interactions among explanatory variables, and survival analysis to trace the trajectory of program terminations over time—the study findings in this section can be summarized as follows:

- Offender characteristics are strongly correlated with the offender’s propensity to comply with program requirements and to complete the 52-week program. Higher risks for failure in the program are associated with offenders who tend to be younger, have limited education, are in a relatively unsettled living condition both economically as well as in terms of family relations, and have an extended history of criminal activities and drug abuse issues. These predictive factors associated with program completion or termination appear consistent with the salient criminogenic factors often cited in the literature for the criminal population in general.<sup>8</sup>
- There is evidence indicating the presence of “program effects,” to the extent that the patterns of both termination and completion exhibit program-level variances, although the nature of these effects remains unspecified and unidentifiable in the models constructed in this study. In fully specified models, the statistical significance associated with program-level effect is substantially reduced. This suggests that the appearance of program effect is to a large extent a reflection of systematic variances of offender characteristics in the programs.
- After controlling for both offender characteristics and program effects, offenders in San Joaquin show a higher propensity for early termination, whereas those in Solano tend to experience a lower rate of termination. Despite their higher risks for early termination, however, San Joaquin offenders are no less likely to complete the program compared with those in other jurisdictions—with the exception of Solano, which shows a higher rate of program completion.

## ***Recidivism***

While compliance with various program requirements is a critical component in the overall batterer intervention strategy, the ultimate goal attached to program completion and other conditions of probation is to hold the offenders accountable for their abusive behavior and to prevent them from re-offense in the future. As a proxy for re-offense, this section relies on re-arrest records from the statewide California State Department of Justice database to examine the patterns of recidivism and risk factors associated with them.

With the arrest records categorized into various offense types, Table 5-E presents re-arrest rates for two specific types: (1) arrests of any charges without differentiating offense types and (2)

arrests in which the charge specifies a spouse/partner as the victim. The length of follow-up time is 12 months from program enrollment until the end of data collection in February 2008.

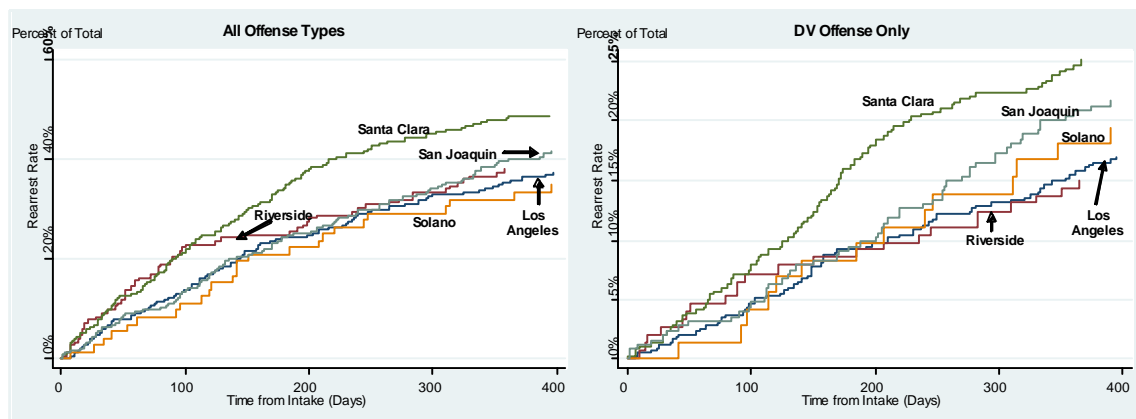
At the level of cross-jurisdictional comparisons alone (i.e., without controlling for confounding factors), Table 5-E shows noticeable differences for both measures of re-arrest. Overall, 40 percent of the study subjects were rearrested within 12 months for various offense types. Re-arrest rates range from 32 percent in Solano to 48 percent in Santa Clara. Of the total re-arrests, approximately one-half are related to domestic violence–specific offenses, leading to an overall domestic violence re-arrest rate of 19 percent.<sup>9</sup> Here the highest re-arrest rate appears again in Santa Clara, at 25 percent. Following behind is San Joaquin at 21 percent, with the remaining jurisdictions between 15 and 20 percent.

**Table 5-E. Re-arrests Within 12 Months After Program Enrollment**

| Jurisdiction | Total Sample | 12-Month Re-arrests for Any Offense |                     | 12-Month Re-arrests for DV Offense |                     |
|--------------|--------------|-------------------------------------|---------------------|------------------------------------|---------------------|
|              |              | Number Re-arrested                  | Percent Re-arrested | Number Re-arrested                 | Percent Re-arrested |
| Los Angeles  | 435          | 155                                 | 36%                 | 69                                 | 16%                 |
| Riverside    | 153          | 58                                  | 38%                 | 23                                 | 15%                 |
| Santa Clara  | 389          | 188                                 | 48%                 | 96                                 | 25%                 |
| Solano       | 72           | 23                                  | 32%                 | 13                                 | 18%                 |
| San Joaquin  | 254          | 102                                 | 40%                 | 53                                 | 21%                 |
| Total        | 1,303        | 526                                 | 40%                 | 254                                | 19%                 |

Note: Without controlling for confounding factors, differences across the jurisdictions are statistically significant at 1% level for both recidivism measures.

**Figure 5-B. Cumulative Re-arrest Rates Within 12 Months After Intake**



The divergent trajectory of re-arrest rates among the jurisdictions can be viewed clearly in Figure 5-B which displays the increasing rates of re-arrests throughout the follow-up period. Santa Clara stands out on both measures in its distinct path compared with other jurisdictions. At around 200 days after intake, San Joaquin’s trajectory for domestic violence re-arrests is also showing signs of moving upward away from Riverside, Los Angeles, and Solano and toward Santa Clara.



It is important to note that, while the overall upward movement of re-arrest rates for all offense types show signs of leveling off over time—increasing from 27 percent at 6 months after intake to 40 percent at 12 months, a rate of change of approximately 50 percent over a follow-up period that doubled in length—there is no sign that the domestic violence re-arrest rates are showing a comparable trend. From an overall rate of 11 percent at 6 months after intake, it maintains a nearly constant rate of increase, almost doubling to 19 percent at 12 months after intake.

## **Descriptive Analysis of Risk Factors**

Given the divergent patterns of re-arrest rates across the jurisdictions, it is necessary to examine the impacts of various risk factors associated with offender characteristics before any independent, system-level impacts can be isolated.

Table 5-F presents 12-month re-arrest rates for all offense types broken out by various categories of risk factors. Without accounting for interaction effects among the factors, differences in re-arrest rates across all the analytical categories are statistically significant.

Table 5-G presents the same breakdowns of risk factors but with re-arrests limited to domestic violence–specific charges. While most risk factors reveal similar correlations with the propensity for domestic violence re-arrests, it is worth noting some important differences between the two re-arrest measures.

The following variables all have statistically significant effects on both re-arrest measures, with offenders exhibiting an *increased* risk of re-offense if:

- The victim is not the offender’s wife;
- The offender has lower educational attainment;
- The offender is African American (compared to other groups);
- The offender speaks English (compared to Hispanics and “others” who do not);
- The offender is younger at the time of intake and younger at the time of his first arrest;
- The offender has an extensive prior arrest history including assaults, domestic violence, and drug-related charges.

Risk factors that are highly correlated with re-arrest rates in general but lacking any discernable impacts on domestic violence re-arrest rates include the following:

- Living arrangements with victim and children;
- Employment status, recent experience of job loss, and income;
- CAGE indicator of drug/alcohol abuse; and
- Presence of felony arrests in prior criminal history.

**Table 5-F. Risk Factors and Re-arrests for All Offense Types 12 Months After Intake**

|                                           | Number of<br>Cases | Percent<br>Arrested in<br>12 Months |                                        | Number of<br>Cases | Percent<br>Arrested in<br>12 Months |
|-------------------------------------------|--------------------|-------------------------------------|----------------------------------------|--------------------|-------------------------------------|
| <u>Live with Victim</u>                   |                    |                                     | <u>Age at Intake</u>                   |                    |                                     |
| No                                        | 769                | 44%                                 | <25                                    | 279                | 51%                                 |
| Yes                                       | 495                | 35%                                 | 25 - 29                                | 261                | 40%                                 |
| Total                                     | 1,264              | 40%                                 | 30 - 39                                | 398                | 38%                                 |
|                                           |                    |                                     | >= 40                                  | 330                | 34%                                 |
| <u>Victim is Wife - Current or Former</u> |                    |                                     | Total                                  | 1,268              | 40%                                 |
| No                                        | 698                | 47%                                 |                                        |                    |                                     |
| Yes                                       | 570                | 31%                                 | <u>CAGE Score (drug/alcohol abuse)</u> |                    |                                     |
| Total                                     | 1,268              | 40%                                 | 0                                      | 466                | 36%                                 |
|                                           |                    |                                     | 1                                      | 158                | 41%                                 |
| <u>Live with Children</u>                 |                    |                                     | 2                                      | 179                | 46%                                 |
| No children                               | 331                | 43%                                 | 3                                      | 137                | 51%                                 |
| Lives with children                       | 426                | 35%                                 | 4                                      | 108                | 53%                                 |
| Visits children regularly                 | 315                | 46%                                 | Total                                  | 1048               | 42%                                 |
| Does not visit children regularly         | 170                | 40%                                 |                                        |                    |                                     |
| Total                                     | 1,242              | 41%                                 | <u>Age at First Arrest</u>             |                    |                                     |
|                                           |                    |                                     | <18                                    | 218                | 53%                                 |
| <u>Lost Job in Past Year</u>              |                    |                                     | 18 - 24                                | 635                | 47%                                 |
| No                                        | 815                | 40%                                 | 25 - 29                                | 166                | 29%                                 |
| Yes                                       | 240                | 49%                                 | 30 - 39                                | 178                | 30%                                 |
| Total                                     | 1055               | 42%                                 | >=40                                   | 104                | 13%                                 |
|                                           |                    |                                     | Total                                  | 1,301              | 40%                                 |
| <u>Employment Status</u>                  |                    |                                     | <u>Total Prior Arrests</u>             |                    |                                     |
| Full-time employed                        | 491                | 33%                                 | 1                                      | 236                | 15%                                 |
| Employed part-time                        | 166                | 46%                                 | 2 - 3                                  | 295                | 32%                                 |
| Not employed                              | 398                | 51%                                 | 4 - 5                                  | 182                | 34%                                 |
| Total                                     | 1,055              | 41%                                 | >=6                                    | 590                | 57%                                 |
|                                           |                    |                                     | Total                                  | 1,303              | 40%                                 |
| <u>Income</u>                             |                    |                                     | <u>Prior Assault Arrests</u>           |                    |                                     |
| \$0 - \$4,999                             | 308                | 42%                                 | 1                                      | 494                | 29%                                 |
| \$5,000 - \$14,999                        | 286                | 48%                                 | 2                                      | 258                | 41%                                 |
| \$15,000 - \$24,999                       | 269                | 39%                                 | 3-4                                    | 275                | 46%                                 |
| \$25,000 - \$39,999                       | 213                | 37%                                 | >=5                                    | 276                | 55%                                 |
| >=\$40,000                                | 109                | 29%                                 | Total                                  | 1,303              | 40%                                 |
| Total                                     | 1,185              | 41%                                 |                                        |                    |                                     |
| <u>Education: Some College or More</u>    |                    |                                     | <u>Prior DV Arrests</u>                |                    |                                     |
| No                                        | 965                | 43%                                 | 1                                      | 694                | 35%                                 |
| Yes                                       | 276                | 29%                                 | 2                                      | 293                | 45%                                 |
| Total                                     | 1,241              | 40%                                 | >=3                                    | 316                | 47%                                 |
|                                           |                    |                                     | Total                                  | 1,303              | 40%                                 |
| <u>Race/Ethnicity</u>                     |                    |                                     | <u>Prior Drug Arrests</u>              |                    |                                     |
| African American                          | 215                | 49%                                 | No                                     | 646                | 32%                                 |
| Hispanic                                  | 667                | 37%                                 | Yes                                    | 657                | 49%                                 |
| Other                                     | 133                | 31%                                 | Total                                  | 1,303              | 40%                                 |
| White                                     | 283                | 46%                                 |                                        |                    |                                     |
| Total                                     | 1,298              | 40%                                 | <u>Prior Felony Arrests</u>            |                    |                                     |
|                                           |                    |                                     | No                                     | 133                | 25%                                 |
| <u>Non-English Speaker</u>                |                    |                                     | Yes                                    | 1,170              | 42%                                 |
| No                                        | 1023               | 43%                                 | Total                                  | 1,303              | 40%                                 |
| Yes                                       | 280                | 30%                                 |                                        |                    |                                     |
| Total                                     | 1,303              | 40%                                 |                                        |                    |                                     |

Note: Differences across comparison groups in rearrest rates are all statistically significant at a minimum of 5% level based on ANOVA.

**Table 5-G. Risk Factors and Re-arrests for Domestic Violence Offenses 12 Months After Intake**

|                                          | Number of Cases | Percent Arrested for DV in 12 Months |                                        | Number of Cases | Percent Arrested for DV in 12 Months |
|------------------------------------------|-----------------|--------------------------------------|----------------------------------------|-----------------|--------------------------------------|
| <u>Live with Victim</u>                  |                 |                                      | <u>Age at Intake**</u>                 |                 |                                      |
| No                                       | 769             | 20%                                  | <25                                    | 279             | 26%                                  |
| Yes                                      | 495             | 19%                                  | 25 - 29                                | 261             | 21%                                  |
| Total                                    | 1,264           | 20%                                  | 30 - 39                                | 398             | 19%                                  |
|                                          |                 |                                      | >= 40                                  | 330             | 14%                                  |
| Victim is Wife - Current or Former**     |                 |                                      | Total                                  | 1,268           | 19%                                  |
| No                                       | 698             | 23%                                  |                                        |                 |                                      |
| Yes                                      | 570             | 15%                                  | <u>CAGE Score (drug/alcohol abuse)</u> |                 |                                      |
| Total                                    | 1,268           | 19%                                  | 0                                      | 466             | 18%                                  |
|                                          |                 |                                      | 1                                      | 158             | 22%                                  |
| <u>Live with Children</u>                |                 |                                      | 2                                      | 179             | 22%                                  |
| No children                              | 331             | 21%                                  | 3                                      | 137             | 23%                                  |
| Lives with children                      | 426             | 19%                                  | 4                                      | 108             | 24%                                  |
| Visits children regularly                | 315             | 21%                                  | Total                                  | 1048            | 20%                                  |
| Does not visit children regularly        | 170             | 14%                                  |                                        |                 |                                      |
| Total                                    | 1,242           | 19%                                  | <u>Age at First Arrest**</u>           |                 |                                      |
|                                          |                 |                                      | <18                                    | 218             | 27%                                  |
| <u>Lost Job in Past Year</u>             |                 |                                      | 18 - 24                                | 635             | 23%                                  |
| No                                       | 815             | 20%                                  | 25 - 29                                | 166             | 11%                                  |
| Yes                                      | 240             | 23%                                  | 30 - 39                                | 178             | 14%                                  |
| Total                                    | 1055            | 20%                                  | >=40                                   | 104             | 8%                                   |
|                                          |                 |                                      | Total                                  | 1,301           | 19%                                  |
| Employment Status                        |                 |                                      |                                        |                 |                                      |
| Full-time employed                       | 491             | 17%                                  | <u>Total Prior Arrests**</u>           |                 |                                      |
| Employed part-time                       | 166             | 23%                                  | 1                                      | 236             | 8%                                   |
| Not employed                             | 398             | 22%                                  | 2 - 3                                  | 295             | 17%                                  |
| Total                                    | 1,055           | 20%                                  | 4 - 5                                  | 182             | 17%                                  |
|                                          |                 |                                      | >=6                                    | 590             | 26%                                  |
| Income                                   |                 |                                      | Total                                  | 1,303           | 19%                                  |
| \$0 - \$4,999                            | 308             | 19%                                  |                                        |                 |                                      |
| \$5,000 - \$14,999                       | 286             | 24%                                  | <u>Prior Assault Arrests**</u>         |                 |                                      |
| \$15,000 - \$24,999                      | 269             | 18%                                  | 1                                      | 494             | 12%                                  |
| \$25,000 - \$39,999                      | 213             | 18%                                  | 2                                      | 258             | 22%                                  |
| >=\$40,000                               | 109             | 15%                                  | 3-4                                    | 275             | 23%                                  |
| Total                                    | 1,185           | 19%                                  | >=5                                    | 276             | 28%                                  |
|                                          |                 |                                      | Total                                  | 1,303           | 19%                                  |
| <u>Education: Some College or More**</u> |                 |                                      |                                        |                 |                                      |
| No                                       | 965             | 22%                                  | <u>Prior DV Arrests**</u>              |                 |                                      |
| Yes                                      | 276             | 12%                                  | 1                                      | 694             | 14%                                  |
| Total                                    | 1,241           | 19%                                  | 2                                      | 293             | 26%                                  |
|                                          |                 |                                      | >=3                                    | 316             | 26%                                  |
| <u>Race/Ethnicity**</u>                  |                 |                                      | Total                                  | 1,303           | 19%                                  |
| African American                         | 215             | 27%                                  |                                        |                 |                                      |
| Hispanic                                 | 667             | 18%                                  | <u>Prior Drug Arrests*</u>             |                 |                                      |
| Other                                    | 133             | 10%                                  | No                                     | 646             | 17%                                  |
| White                                    | 283             | 21%                                  | Yes                                    | 657             | 22%                                  |
| Total                                    | 1,298           | 19%                                  | Total                                  | 1,303           | 19%                                  |
|                                          |                 |                                      |                                        |                 |                                      |
| <u>Non-English Speaker**</u>             |                 |                                      | <u>Prior Felony Arrests</u>            |                 |                                      |
| No                                       | 1023            | 21%                                  | No                                     | 133             | 15%                                  |
| Yes                                      | 280             | 13%                                  | Yes                                    | 1,170           | 20%                                  |
| Total                                    | 1,303           | 19%                                  | Total                                  | 1,303           | 19%                                  |

\* Statistically significant at 5% level; \*\* statistically significant at 1 % level.

The fact that many of these non-significant factors capture some important aspects of the offender's socioeconomic status, which under most circumstances are strongly correlated with criminality, suggests the existence of a different set of factors contributing to batterer behavior in a criminal population. The persistence in the rate of domestic violence re-offense as shown in the survival chart (Table 5-B) above appears to point to the same conclusion. While the offender population in this study exhibits many of the characteristics that are typical of chronic offenders generally, their domestic violence may involve a different set of complex factors that are quite distinct from the general criminal population.

## **Regression Analysis of Recidivism**

Having examined the different patterns of re-arrest rates across the jurisdictions as well as their correlations with various risk factors, in this section we apply the same analytical strategy employed in the analysis of program completion and termination to identify system-level impacts on recidivism.

### **No Clear Evidence of Program Effects on Re-arrest**

Regression results for 12-month recidivism are displayed in Table 5-H for re-arrests for all offense types and Table 5-I for domestic violence re-arrests. At the level of system impacts (i.e., program effects and overall jurisdictional effects), analysis results for re-arrests are especially interesting in contrast to the findings on program completion and termination. In the previous section we saw that BIPs appear to exert some systematic impacts on offenders' propensity for program termination and completion, even though the precise impact is unspecified.

In contrast to the effect of programs on completion and termination, there is no clear evidence from various regression models indicating the presence of strong program effects on the likelihood of re-arrests. The lack of a program effect on re-arrest holds for all offense types and for domestic violence offenses. This finding can be seen first by comparing regression models 1 and 2. This pair of models allows for an examination of correlation patterns of model *error terms*, which provide a measure of the underlying robustness and bias of the regression model. Santa Clara shows a statistically significant impact in models 1 and 2 on both measures of re-arrests and, in both cases, is associated with *higher* risks of re-arrest than in other jurisdictions. There is almost no change between the two models in the level of statistical significance attached to these estimated effects. Within the framework of ordinary regression models, this result indicates the underlying stability of the regression error terms unaffected by any *cluster* patterns that might exist across programs.

Further evidence regarding the absence of program effects on re-arrests can be found in the analysis results from multilevel models, in which program-level effects are explicitly modeled to measure the size of their variances as well as the statistical significance of the variances. While differing in terms of the control variables specified in the models, estimates of program-level variance with respect to re-arrest patterns do not produce any significant results in models 5 through 7.

## Jurisdictional Effects on Re-arrest

Without statistical evidence of a link between an offender's propensity for re-arrest and the program in which he was enrolled, jurisdictional effects that remain in the model can be attributed to either one of two sources. Jurisdictional effects might still be the result of confounding effects of observable offender profile variables or, after controlling for these individual-level variables, jurisdictional effects may in fact be a result of particular system-level organizational characteristics. In contrast to the analysis results for program completion and termination, the various regression models for re-arrests provide evidence of different patterns of jurisdictional impacts on re-arrest rates.

The regression models for program completion and termination in the previous section show a complex relationship between system impacts on completion and termination. System impacts not only vary by the specific outcome measure examined, but the stability of estimated effects—to the extent the effects are statistically significant at the jurisdictional level—appears to be subject to variations in model specifications. Thus estimated system impacts, as well as the statistical significance for the estimates, may increase or decrease depending on the control variables included in the models and on the specification of a multilevel structure for the models.

Compared with these relatively unstable findings associated with system impacts on program completion and termination, regression results in Table 5-H and Table 5-I provide evidence of statistically significant and stable impacts in Santa Clara in its jurisdictional effect on re-arrest rates. While mediated to some extent by offender profile variables included in different models, offenders in Santa Clara have consistently and statistically significant *higher* risks for re-arrest within 12 months after intake. This finding is consistent for re-arrest for all offense types as well as re-arrest only for domestic violence offenses.

In addition to the jurisdictional effect that we see in Santa Clara across both measures of re-arrest, offenders in Solano have a statistically significant *lower* risk of re-arrest but only on the measure of re-arrests for all offense types and not domestic violence re-offenses. It should be noted also that the evidence of jurisdictional impact in Solano is less robust than the evidence for Santa Clara in that it emerges only after differences in offender profiles have been controlled for in the models.

**Table 5-H. Regression Results for 12-Month Re-arrests for All Offense Types**

| Dependent Variable:<br>12-Month Rearrests of All Offense Types                             | Logistic Regression Models |                   |                   |                   | Multilevel Logistic Regression Models |                   |                   |
|--------------------------------------------------------------------------------------------|----------------------------|-------------------|-------------------|-------------------|---------------------------------------|-------------------|-------------------|
|                                                                                            | (1)                        | (2)               | (3)               | (4)               | (5)                                   | (6)               | (7)               |
| <i>Jurisdiction (Los Angeles as the base comparison group)</i>                             |                            |                   |                   |                   |                                       |                   |                   |
| Riverside                                                                                  | 1.103<br>(0.50)            | 1.103<br>(0.56)   | 0.938<br>(0.33)   | 1.112<br>(0.39)   | 1.110<br>(0.48)                       | 0.941<br>(0.28)   | 1.131<br>(0.47)   |
| Santa Clara                                                                                | 1.690<br>(3.68)**          | 1.690<br>(3.04)** | 1.655<br>(2.89)** | 1.683<br>(2.73)** | 1.625<br>(2.88)**                     | 1.635<br>(2.98)** | 1.650<br>(2.40)*  |
| Solano                                                                                     | 0.848<br>(0.61)            | 0.848<br>(0.60)   | 0.580<br>(2.10)*  | 0.335<br>(2.93)** | 0.830<br>(0.63)                       | 0.577<br>(1.63)   | 0.332<br>(2.08)*  |
| San Joaquin                                                                                | 1.212<br>(1.18)            | 1.212<br>(0.94)   | 0.943<br>(0.38)   | 0.865<br>(0.84)   | 1.185<br>(0.89)                       | 0.941<br>(0.32)   | 0.858<br>(0.65)   |
| <i>No children as base comparison group</i>                                                |                            |                   |                   |                   |                                       |                   |                   |
| Lives with children                                                                        |                            |                   | 0.889<br>(0.61)   | 0.929<br>(0.35)   |                                       | 0.889<br>(0.70)   | 0.928<br>(0.37)   |
| Visits children regularly                                                                  |                            |                   | 1.266<br>(1.46)   | 1.240<br>(1.09)   |                                       | 1.263<br>(1.38)   | 1.236<br>(1.05)   |
| Does not visit children regularly                                                          |                            |                   | 0.924<br>(0.35)   | 1.015<br>(0.06)   |                                       | 0.922<br>(0.39)   | 1.013<br>(0.05)   |
| Education: some college or more                                                            |                            |                   | 0.474<br>(4.90)** | 0.632<br>(2.16)*  |                                       | 0.470<br>(4.68)** | 0.627<br>(2.45)*  |
| Victim is wife--former and current                                                         |                            |                   | 0.548<br>(4.65)** | 0.731<br>(2.06)*  |                                       | 0.549<br>(4.55)** | 0.728<br>(1.92)   |
| Non-English speaker                                                                        |                            |                   | 0.613<br>(3.34)** | 1.463<br>(1.75)   |                                       | 0.618<br>(2.92)** | 1.477<br>(1.71)   |
| <i>Employment Status (employed full-time as base comparison group)</i>                     |                            |                   |                   |                   |                                       |                   |                   |
| Employed part-time                                                                         |                            |                   |                   | 1.335<br>(1.15)   |                                       |                   | 1.337<br>(1.34)   |
| Not employed                                                                               |                            |                   |                   | 1.481<br>(2.43)*  |                                       |                   | 1.481<br>(2.29)*  |
| <i>Race/ethnicity (African American as base comparison group)</i>                          |                            |                   |                   |                   |                                       |                   |                   |
| Hispanic                                                                                   |                            |                   |                   | 0.621<br>(2.20)*  |                                       |                   | 0.616<br>(2.16)*  |
| Other                                                                                      |                            |                   |                   | 0.984<br>(0.04)   |                                       |                   | 0.978<br>(0.06)   |
| White                                                                                      |                            |                   |                   | 0.885<br>(0.47)   |                                       |                   | 0.872<br>(0.56)   |
| <i>Total prior arrests of any offense (single, first offense as base comparison group)</i> |                            |                   |                   |                   |                                       |                   |                   |
| 2-3                                                                                        |                            |                   |                   | 3.238<br>(4.36)** |                                       |                   | 3.259<br>(3.84)** |
| 4-5                                                                                        |                            |                   |                   | 2.921<br>(2.94)** |                                       |                   | 2.945<br>(3.04)** |
| >=6                                                                                        |                            |                   |                   | 8.309<br>(5.62)** |                                       |                   | 8.338<br>(5.60)** |
| <i>Total prior arrests of DV offenses (single, first offense as base comparison group)</i> |                            |                   |                   |                   |                                       |                   |                   |
| 2                                                                                          |                            |                   |                   | 0.934<br>(0.36)   |                                       |                   | 0.938<br>(0.33)   |
| >=3                                                                                        |                            |                   |                   | 0.651<br>(1.94)   |                                       |                   | 0.653<br>(2.09)*  |
| Had prior felony arrests                                                                   |                            |                   |                   | 1.428<br>(1.30)   |                                       |                   | 1.419<br>(1.16)   |
| Had prior arrests for drug offenses                                                        |                            |                   |                   | 0.886<br>(0.70)   |                                       |                   | 0.883<br>(0.69)   |
| Age at intake                                                                              |                            |                   |                   | 0.972<br>(2.48)*  |                                       |                   | 0.972<br>(2.46)*  |
| Age at first arrest                                                                        |                            |                   |                   | 0.990<br>(0.70)   |                                       |                   | 0.990<br>(0.69)   |
| <i>CAGE Score (0 as base comparison group)</i>                                             |                            |                   |                   |                   |                                       |                   |                   |
| CAGE = 1                                                                                   |                            |                   |                   | 1.005<br>(0.02)   |                                       |                   | 1.014<br>(0.06)   |
| CAGE = 2                                                                                   |                            |                   |                   | 1.572<br>(2.33)*  |                                       |                   | 1.576<br>(2.14)*  |
| CAGE = 3                                                                                   |                            |                   |                   | 1.301<br>(1.03)   |                                       |                   | 1.312<br>(1.15)   |
| CAGE = 4                                                                                   |                            |                   |                   | 1.239<br>(0.88)   |                                       |                   | 1.244<br>(0.86)   |
| Program level variance                                                                     |                            |                   |                   |                   | 0.000<br>(0.00)                       | 0.000<br>(0.00)   | 0.000<br>(0.00)   |
| Observations                                                                               | 1,303                      | 1,303             | 1,202             | 941               | 1,303                                 | 1,202             | 941               |

Absolute value of z statistics in parentheses; \* significant at 5%; \*\* significant at 1%.

† In models 2 to 4, Huber-White sandwich estimate of variance is used for robust estimates of standard errors that adjust for intraclass correlations at the level of batterer intervention programs.

**Table 5-I. Regression Results for 12-Month Re-arrests for Domestic Violence Offenses**

| Dependent Variable:<br>12-Month Rearrests of DV Offenses                                   | Logistic Regression Models |                   |                   |                   | Multilevel Logistic Regression Models |                   |                   |
|--------------------------------------------------------------------------------------------|----------------------------|-------------------|-------------------|-------------------|---------------------------------------|-------------------|-------------------|
|                                                                                            | (1)                        | (2)               | (3)               | (4)               | (5)                                   | (6)               | (7)               |
| <i>Jurisdiction (Los Angeles as the base comparison group)</i>                             |                            |                   |                   |                   |                                       |                   |                   |
| Riverside                                                                                  | 0.938<br>(0.24)            | 0.938<br>(0.31)   | 0.812<br>(1.08)   | 0.854<br>(0.85)   | 0.938<br>(0.24)                       | 0.812<br>(0.76)   | 0.854<br>(0.52)   |
| Santa Clara                                                                                | 1.738<br>(3.14)**          | 1.738<br>(3.25)** | 1.627<br>(2.93)** | 1.558<br>(2.48)*  | 1.738<br>(3.14)**                     | 1.627<br>(2.57)*  | 1.558<br>(1.96)*  |
| Solano                                                                                     | 1.169<br>(0.47)            | 1.169<br>(0.49)   | 0.856<br>(0.63)   | 0.484<br>(1.69)   | 1.169<br>(0.47)                       | 0.856<br>(0.39)   | 0.484<br>(1.24)   |
| San Joaquin                                                                                | 1.399<br>(1.66)            | 1.399<br>(1.54)   | 1.080<br>(0.40)   | 0.948<br>(0.25)   | 1.399<br>(1.66)                       | 1.080<br>(0.35)   | 0.948<br>(0.21)   |
| <i>No children as base comparison group</i>                                                |                            |                   |                   |                   |                                       |                   |                   |
| Lives with children                                                                        |                            |                   | 1.059<br>(0.29)   | 1.025<br>(0.11)   |                                       | 1.059<br>(0.29)   | 1.025<br>(0.11)   |
| Visits children regularly                                                                  |                            |                   | 1.014<br>(0.07)   | 0.961<br>(0.17)   |                                       | 1.014<br>(0.07)   | 0.961<br>(0.17)   |
| Does not visit children regularly                                                          |                            |                   | 0.546<br>(2.57)*  | 0.495<br>(2.61)** |                                       | 0.546<br>(2.17)*  | 0.495<br>(2.10)*  |
| Education: some college or more                                                            |                            |                   | 0.412<br>(4.04)** | 0.510<br>(2.59)** |                                       | 0.412<br>(4.11)** | 0.510<br>(2.81)** |
| Victim is wife - former and current                                                        |                            |                   | 0.621<br>(3.11)** | 0.869<br>(0.73)   |                                       | 0.621<br>(2.88)** | 0.869<br>(0.72)   |
| Non-English Speaker                                                                        |                            |                   | 0.556<br>(3.80)** | 0.935<br>(0.31)   |                                       | 0.556<br>(2.76)** | 0.935<br>(0.24)   |
| <i>Employment Status (employed full-time as base comparison group)</i>                     |                            |                   |                   |                   |                                       |                   |                   |
| Employed part-time                                                                         |                            |                   |                   | 1.079<br>(0.32)   |                                       |                   | 1.079<br>(0.30)   |
| Not employed                                                                               |                            |                   |                   | 0.981<br>(0.11)   |                                       |                   | 0.981<br>(0.10)   |
| <i>Race/ethnicity (African American as base comparison group)</i>                          |                            |                   |                   |                   |                                       |                   |                   |
| Hispanic                                                                                   |                            |                   |                   | 0.553<br>(2.78)** |                                       |                   | 0.553<br>(2.36)*  |
| Other                                                                                      |                            |                   |                   | 0.460<br>(1.85)   |                                       |                   | 0.460<br>(1.84)   |
| White                                                                                      |                            |                   |                   | 0.689<br>(1.59)   |                                       |                   | 0.689<br>(1.36)   |
| <i>Total prior arrests of any offense (single, first offense as base comparison group)</i> |                            |                   |                   |                   |                                       |                   |                   |
| 2-3                                                                                        |                            |                   |                   | 2.647<br>(2.76)** |                                       |                   | 2.647<br>(2.49)*  |
| 4-5                                                                                        |                            |                   |                   | 2.194<br>(1.68)   |                                       |                   | 2.194<br>(1.75)   |
| >=6                                                                                        |                            |                   |                   | 3.700<br>(2.30)*  |                                       |                   | 3.700<br>(2.85)** |
| <i>Total prior arrests of DV offenses (single, first offense as base comparison group)</i> |                            |                   |                   |                   |                                       |                   |                   |
| 2                                                                                          |                            |                   |                   | 1.526<br>(1.79)   |                                       |                   | 1.526<br>(1.91)   |
| >=3                                                                                        |                            |                   |                   | 1.418<br>(1.28)   |                                       |                   | 1.418<br>(1.49)   |
| Had prior felony arrests                                                                   |                            |                   |                   | 1.013<br>(0.04)   |                                       |                   | 1.013<br>(0.04)   |
| Had prior arrests for drug offenses                                                        |                            |                   |                   | 0.802<br>(1.26)   |                                       |                   | 0.802<br>(1.08)   |
| Age at intake                                                                              |                            |                   |                   | 0.961<br>(2.50)*  |                                       |                   | 0.961<br>(2.89)** |
| Age at first arrest                                                                        |                            |                   |                   | 1.011<br>(0.54)   |                                       |                   | 1.011<br>(0.59)   |
| <i>CAGE Score (0 as base comparison group)</i>                                             |                            |                   |                   |                   |                                       |                   |                   |
| CAGE = 1                                                                                   |                            |                   |                   | 1.234<br>(0.82)   |                                       |                   | 1.234<br>(0.83)   |
| CAGE = 2                                                                                   |                            |                   |                   | 1.511<br>(1.61)   |                                       |                   | 1.511<br>(1.68)   |
| CAGE = 3                                                                                   |                            |                   |                   | 1.152<br>(0.55)   |                                       |                   | 1.152<br>(0.51)   |
| CAGE = 4                                                                                   |                            |                   |                   | 1.297<br>(1.10)   |                                       |                   | 1.297<br>(0.89)   |
| Program level variance                                                                     |                            |                   |                   |                   | 0.000<br>(0.00)                       | 0.000<br>(0.00)   | 0.000<br>(0.00)   |
| Observations                                                                               | 1,303                      | 1,303             | 1,202             | 941               | 1,303                                 | 1,202             | 941               |

Absolute value of z statistics in parentheses; \* significant at 5%; \*\* significant at 1%.

<sup>†</sup> In models 2 to 4, Huber-White sandwich estimate of variance is used for robust estimates of standard errors that adjust for intraclass correlations at the level of batterer intervention programs.

In addition to the presence or absence of impacts at the program and jurisdiction level, regression results for offender profile variables included in the various regression models reveal patterns fairly similar to those presented in the descriptive analysis of individual risk factors in the previous section. In general, these variables are better predictors of re-arrests for undifferentiated offense types than as predictors of re-arrests for domestic violence offenses. Variables with a statistically significant impact on both re-arrest measures include education, race/ethnicity, total prior arrest history, and age at intake. Specifically, offenders in the sample have a lower propensity for re-arrest if they have some college education or more; are of Hispanic descent; have fewer prior arrests; and are older at the time of intake.

Risk factors that are significantly associated with re-arrests in general but lacking strong correlations with domestic violence re-offenses include offender relationship with the victim, employment status, and CAGE indicator of drug/alcohol abuse. It is worth noting that the variables that are the best predictors of re-arrests for all offense types rather than domestic violence offenses are also variables that exert a greater impact on termination than on completion in the regression models examining program termination and completion. This suggests the possibility that early termination from the programs and the overall propensity for re-arrests could be driven by a similar set of risk factors.

Many variables that did not show any significant impact on program completion or termination previously have similar results in the regression models for re-arrests. These include prior arrests on felony charges, prior arrests on drug/alcohol charges, and age at first arrest. The same reason presented above explaining their lack of statistical significance— i.e., the effect of a “weaker” variable being diminished when correlated with other variables with more robust effects— appears to be applicable here as well.

In addition to the variables examined above showing different impacts on the two re-arrest measures, there are two variables in the models that produced somewhat unexpected results: prior arrest history of domestic violence charges and offender’s relationship with children. For prior arrests on domestic violence charges, the correlation of that variable with overall prior arrest history is expected to reduce its predictive power in the models—or when its predictive power is unaffected by other correlated variables, it is expected to reveal stronger association with re-arrests for domestic violence offense than with the overall, undifferentiated re-arrest measure. The analysis results show that, compared with offenders who have only one prior arrest for domestic violence charges, those with three or more domestic violence prior arrests have a *lower* propensity for re-arrest of any offense type, other factors being controlled for. With regard to domestic violence re-arrests, there is a marginally significant impact (less than 90 percent confidence level) associated with *increased* risks for those with two prior arrests on domestic violence charges.

Previous analysis of an offender’s relationship with his children and program termination or completion suggests that compared with childless offenders, those who have children and are living with them have a better chance of completing the program (with marginal statistical significance) and a lower chance of being terminated from the program (with substantially larger statistical significance). In fully specified models, however, the statistical impact of this variable



is almost completely overtaken by other relevant socioeconomic variables with more robust effects. In regression models for re-arrest on domestic violence charges, offenders who have children but maintain irregular contact with them reveal consistently lower risks of domestic violence re-arrest than offenders who don't have children. With this factor showing persistent effects in various models specified with different control variables, it is not clear as to the nature this connection. In contrast to its impact on domestic violence re-arrests, an offender's relationship with his children reveals no statistical connection to the overall propensity for re-arrests.

## **Summary of Findings**

To summarize the findings presented above regarding jurisdictional and program impacts on program completion, program termination, and recidivism, the following conclusions can be drawn:

- Batterer intervention programs tend to exert an independent influence on an offender's probability of maintaining continuous attendance (allowing for absences of varying degrees) in the program and ultimately successfully completing the 52-week program.
- When variances in offender characteristics across the programs are controlled for, program effects on both termination and completion are reduced substantially;
- While variances in completion and termination across the programs reflect different characteristics of the offenders who tend to enroll in specific programs, there is no evidence indicating any systematic variance across the programs in their clients' propensity for re-arrest.
- After accounting for individual- and program-level variances, jurisdictional differences remain persistent in both program outcomes and re-arrest rates.
  - Looking at system-level variance in program outcomes, offenders in San Joaquin exhibit higher risks for termination. Completion rates for offenders in San Joaquin, however, are no worse than those in other jurisdictions with the exception of Solano. In Solano, offenders in the sample are less likely to fail in the program with an early termination and more likely to continue through the end to complete the 52-week program;
  - Looking at system-level variance in re-arrest rates across the jurisdictions, offenders in Santa Clara show a persistently higher risk of re-offense while offenders in the sample in Solano showed a lower risk for re-offense.

While the different levels of statistical analysis presented clear answers on many questions, in particular with regard to the presence or absence of system impacts at the level of programs and jurisdictions, many questions remain unanswered. For example, the higher risks for re-arrests associated with offenders in Santa Clara are clearly shown in the regression models; they cannot be attributed to differences in either offender profiles or program characteristics in Santa Clara compared with other jurisdiction, as these potential confounding factors have been incorporated into the analysis.

Santa Clara is generally known for its close coordination among the justice system partners; however, this analysis does not provide an answer about the specific components of the system

or the causal mechanisms through which the particular intervention strategy adopted in the jurisdiction has led to higher re-arrest rates.

We should also point out that, from an analytical, modeling perspective, statistical findings are inherently constrained by the adequacy of the models in terms of observable confounding factors included in the models as controls as well as the underlying validity and accuracy of measurements for these control variables. To the extent measurement problems that might exist in some variables (such as the accuracy of the CAGE scores, or the potential undercounting of prior arrest records as a true measure of *actual* incidents of criminal activities) are randomly distributed across the various analytical categories, they should not lead to biased results within the regression analysis framework.

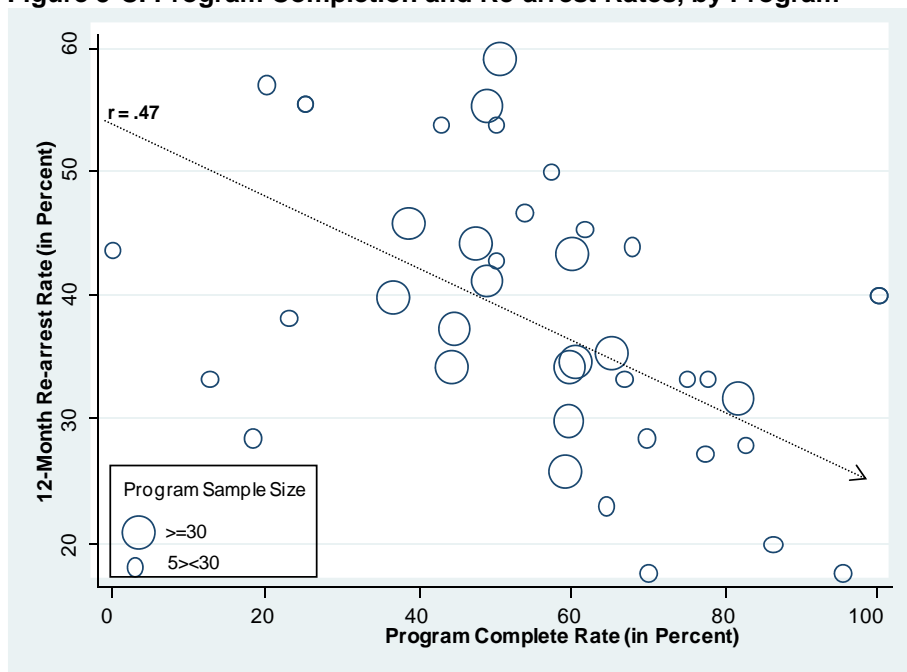
The bigger issue in a quasi-experimental design such as this study has to do with the adequacy of the models in fully capturing the factors that might affect the outcome variables examined in various models—the problem of omitted variables or underspecified models.<sup>10</sup> We pointed out in the analysis of offender profiles that offenders in Santa Clara—and similarly those in San Joaquin—tend to exhibit greater prevalence in various risk factors often associated with chronic offenders, whether measured by socio-demographic and economic factors, patterns of family relations, or criminal history variables. While these control variables are highly predictive of the various outcome variables examined, thus *explaining away* some variances associated with jurisdictional effects, a substantial portion of the outcome variances remains *unexplained* owing to relevant variables that are unavailable, measured with questionable validity (such as Conflict Tactics Scale 2, or CTS2, scores gathered for this study) or simply unobservable. It is therefore plausible to assume that, given a more fully specified model consisting of more pertinent, predictive variables in relation to offenders' propensity for re-offense behavior, Santa Clara's higher re-arrest rate could be further explained away by confounding factors currently not controlled for in the models.

### ***Relationship Between Program Completion and Recidivism***

Having discussed the methodological limitation of the study, and in the absence of additional explanatory variables to augment the predictive models, we turn here to an examination of the relationship between program completion and re-arrest rates. This evaluation may shed some light on the different patterns of interactions among different system players in their efforts to hold batterers accountable.

One major challenge of modeling the relationship between program completion and re-arrest is the fact that these two events do not always sequence consecutively. In other words, re-arrest might occur before or after program completion and may or may not be directly linked to one another. As a result, any interaction that might exist between the two is not amenable to a clearly delineated causal model. At the aggregate level, however, one can examine completion and re-arrest rates by program.

**Figure 5-C. Program Completion and Re-arrest Rates, by Program**

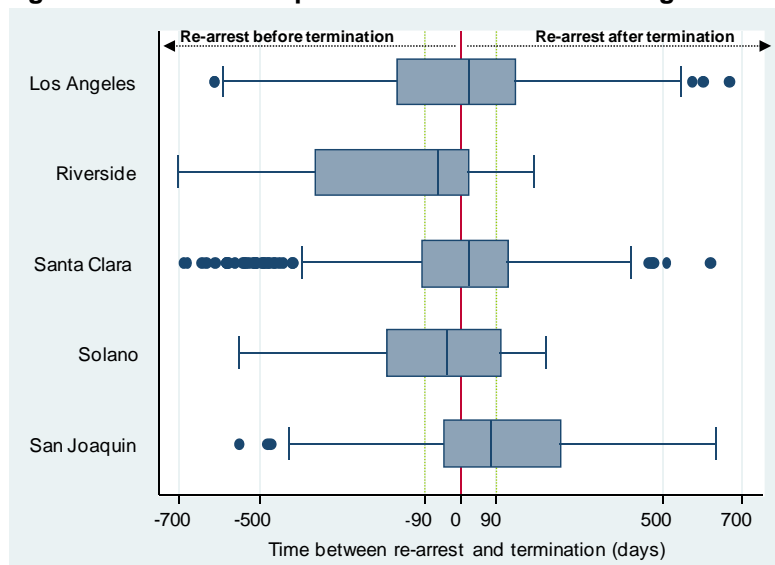


In Figure 5-C, each data point represents a batterer intervention program, with program sample size differentiated by the size of the circle. The horizontal axis represents program completion rate and the vertical axis represents re-arrest rate for any offense within 12 months after intake, with both rates measured at the aggregate program level.

The dispersion pattern of the data points, clustered loosely around the downward-sloping line, indicates the existence of some correlation between the two outcome measures, valid for both large and small programs. The inverse relationship means that a program with a lower re-arrest rate (moving downward on the vertical axis toward zero) is likely to see a larger percentage of its clients completing the program (moving farther to the right on the horizontal axis). The correlation is a modest one—the correlation coefficient of .47 between the two indicates that approximately 25 percent of the change in one variable can be explained by change in the other variable. However, this correlation suggests a series of causal chains, from offender’s risk factors to propensity for re-arrest and other forms of probation violation to program termination and program completion, that constrains the performance of a program.

To be sure, the causal chain represents merely a probabilistic model of how different factors might be connected to one another under normal circumstances. In response to the observed causal connections—subject to different interpretations of their meaning—different strategies, policies, and practices at different levels of the system nevertheless have sufficient room for discretionary actions to intervene and shape the interaction among different factors. An analysis of the timing of re-arrests in relation to program termination may provide some insights about the interaction between these two events as well as the different actions taken by the justice system partners in response to re-offenses committed by the offenders.

**Figure 5-D. Relationship Between Re-arrest and Program Termination**



The box-whisker plots in Figure 5-D display the timing of re-arrests in relation to the timing of program termination, which is represented by the vertical line at zero (days). To the left of that vertical line are re-arrests prior to termination from the program, and to the right are re-arrests after termination.

Re-arrests prior to termination, particularly those shortly before termination, may constitute causes for termination. In a different scenario, an offender may fail to show up for multiple consecutive sessions, which could lead to his termination from the program, following which he may reoffend. While the termination and re-arrest described in this scenario are related, they are certainly not linked to one another in a narrow causal sense; cases falling into this scenario are represented to the right of zero days.

With 50 percent of the cases in each jurisdiction represented by the box width, there appear to be significant variances in the timing of re-arrest and termination. The dispersion patterns in Riverside and San Joaquin present a mirror image of each other, with the majority of cases in Riverside (68 percent) rearrested before termination from program, and almost the same proportion in San Joaquin (67 percent) rearrested following termination. In Santa Clara and Los Angeles, re-arrests are fairly evenly distributed in relation to program termination with narrow majorities (56 and 58 percent respectively) of those who were re-arrested having already been terminated from program. In contrast, in Solano a majority (59 percent) of offenders who were re-arrested had not yet been terminated from program.

It is not clear how (or even whether) the different patterns depicted in Figure 5-D have any bearing on the overall jurisdictional effect discussed above. The varying prevalence of post-termination re-arrests may result directly from the different forms of sanctions implemented in individual jurisdictions, varying in the efficacy of their deterrence effect in holding offenders accountable. These patterns appear to point to system-level differences that may not be captured in the regression analysis of outcomes.

## Changes in Beliefs and Attitudes

The outcome measures examined so far, including program completion and termination as well as re-arrest, are proxy measures of behavioral changes critical to the intervention strategy for domestic violence offenders. According to the logic model described in Chapter 1, Figure 1-A, behavioral changes resulting from system interventions and program participation may at some level be accompanied (or preceded) by psychosocial changes in offenders' beliefs and attitudes. Theoretically, psychosocial changes as measured by the pre- and post-BIP Process Survey could be incorporated into the full model to examine the causal path connecting program effects to psychosocial changes and to behavior changes.

As noted in Chapter 1, however, there are substantial gaps in the data collected for this instrument. When pre- and post-survey responses are matched by individual offenders at the final stage of data collection, only about 15 percent (233 offenders) of the entire sample have complete data available for measuring pre-post changes. Because of the small sample size as well as the potential sample selection bias, limited analysis of the data is presented, with the results shown in Table 5-J; a graphic display of the different patterns of the pre-post changes for the five subscales appears in Figure 5-E.

**Table 5-J. Pre-Post Changes of BIP Process Survey**

|             | Sample Size | Personal Responsibility |      |         | Power and Control |      |         | Understanding the Effect of Abuse on Others |      |         | Dependency |      |         | Anger Management |      |         |
|-------------|-------------|-------------------------|------|---------|-------------------|------|---------|---------------------------------------------|------|---------|------------|------|---------|------------------|------|---------|
|             |             | Pre                     | Post | Change  | Pre               | Post | Change  | Pre                                         | Post | Change  | Pre        | Post | Change  | Pre              | Post | Change  |
| Los Angeles | 74          | 4.59                    | 4.72 | 0.13    | 4.67              | 4.78 | 0.12    | 3.12                                        | 3.34 | 0.23    | 4.39       | 4.55 | 0.16    | 4.43             | 4.60 | 0.16 *  |
| Riverside   | 30          | 4.28                    | 4.59 | 0.31 *  | 4.66              | 4.82 | 0.16    | 3.28                                        | 3.40 | 0.12    | 3.99       | 4.24 | 0.26    | 4.46             | 4.54 | 0.09    |
| Santa Clara | 90          | 4.89                    | 5.19 | 0.30 ** | 4.49              | 4.84 | 0.35 ** | 3.78                                        | 4.42 | 0.64 ** | 4.27       | 4.53 | 0.25 ** | 4.68             | 4.79 | 0.12 *  |
| San Joaquin | 32          | 4.56                    | 4.87 | 0.31 ** | 4.62              | 4.76 | 0.14    | 3.28                                        | 3.59 | 0.31 *  | 4.32       | 4.52 | 0.20 *  | 4.45             | 4.68 | 0.23 *  |
| Total       | 226         | 4.68                    | 4.91 | 0.24 ** | 4.59              | 4.80 | 0.20 ** | 3.41                                        | 3.79 | 0.38 ** | 4.29       | 4.50 | 0.21 ** | 4.55             | 4.68 | 0.14 ** |

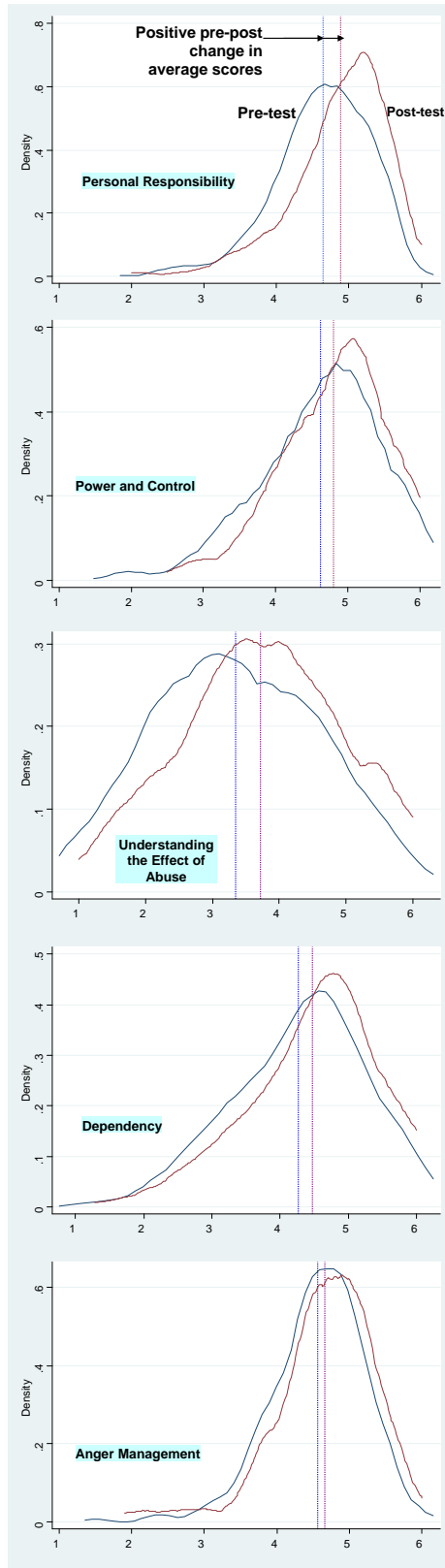
Note: BIP Process Survey not administered in Solano County due to small overall sample size.

\* Statistically significant at 5% level, \*\* statistically significant at 1 % level, based on one-tailed t-test.

Ranging in values from 1 to 6, Table 5-J shows the five subscale scores from both the pre- and post-surveys, as well as the differences between the pre- and post-scores; due to its small sample size (seven cases with available data), results for Solano are not shown in the table. As higher scores associated with each subscale are deemed more desirable, positive changes measured by the instrument would lead to higher scores in the post-survey results. With varying degrees of statistical significance—some inevitably affected by small sample sizes, pre-post differences reveal positive changes, *on average*, in all five subscales and across all jurisdictions.

Overall, the subscale measuring the offenders' capacity for understanding the negative impact of their abusive behavior on others—closely related to the concept of empathy—shows the largest positive change, in particular for offenders in Santa Clara, where the average increases from 3.8 to 4.4. On the other hand, the anger management subscale, one of the topics covered commonly in the BIP curricula, shows relatively small positive changes. The other three subscales—measures of personal responsibility, power and control, and dependency—show similar levels of modest changes.

**Figure 5-E. Pre- and Post-BIP Process Survey**



In connection with the relative size of pre-post changes for the five subscales, Figure 5-E further shows differences in their underlying distribution patterns. It appears that three subscales—personal responsibility, power and control, and dependency—share similar characteristics, with the distribution curves skewed to the left. The empathy subscale (understanding the effect of abuse) approximates a normal curve in its distribution. Anger management also displays a normal curve but with a much narrower dispersion, suggesting greater homogeneity in the responses compared with the other subscales.

Comparisons across jurisdictions show no statistically significant differences for any of the subscales except for the empathy subscale, where respondents in Santa Clara appear to show significant change relative to other jurisdictions.

## **Summary**

In this chapter we examined offender outcomes in terms of two principal measures: program completion and re-arrest. Attendance records for each offender enrolled in the study were analyzed to discern patterns in attendance, absences, and termination. We also identified offender characteristics that were strongly correlated with program termination and completion. Those risk factors were used as control variables in various regression analyses that were used to try to answer the central questions of the study: whether system impacts vary significantly across the jurisdictions; whether the impacts vary systematically across BIPs within a jurisdiction; and, whether program level variance accounts for differences in jurisdictional effects. Finally, we evaluated the findings of the BIP process survey which was administered to study enrollees at intake and just prior to program completion to attempt to measure psychosocial changes in offenders resulting from program enrollment. The next chapter summarizes the major findings of this study and discusses the implications of those findings for policy and future research.

## Endnotes Chapter 5

---

- 1 See *Keeping the Promise: Victim Safety and Batterer Accountability*, Report to the California Attorney General from the Task Force on Local Criminal Justice Response to Domestic Violence (June 2005), p. 63; and *Batterer Intervention Programs: County Probation Departments Could Improve Their Compliance with State Law, but Progress in Batterer Accountability Also Depends on the Courts*, California State Auditor, Bureau of State Audits (November 2006), pp. 22–24.
2. Since a case would not be included in the sample for the calculation of completion rate if updated records indicated that the offender was still in the program, updated information available from programs, to the extent it varies among them, would likely affect the validity of the analysis. Therefore, an analysis of program intake and completion dates was conducted to examine the length of follow-up time across the jurisdictions. No evidence was found that suggests the existence of systematic sampling biases that would threaten the validity of program completion comparisons.
3. More specifically, the estimation problem concerns the error terms in regression models. In classical linear regression models, when error terms are correlated systematically—in the current context, across individuals within the same program—the assumption of constant variance of the error term is violated, leading to biased estimates of standard errors.
4. In testing and constructing various models, the process involves selecting control variables to assess their interactions (collinearity) as well as determining the appropriate forms of their measurements, such as transformation of a continuous scale to categories to evaluate the existence of nonlinear, threshold effects. Stability of the models is also assessed in response to the different patterns of missing data, depending on different sets of variables included in the models. Appendix L provides regression results for the same models discussed in this chapter, but in each table the analysis samples are restricted to the same sample sizes. Sample size restrictions changed to different degrees the coefficient estimates and significance levels for some variables. They do not, however, affect the substantive findings of the study.
5. Generalized Linear Latent and Mixed Models (GLLAMM) implemented in Stata, a program created by Sophia Rabe-Hesketh. See S. Rabe-Hesketh and A. Skrondal, *Multilevel and Longitudinal Modeling Using Stata* (2008), Stata Press.
6. It is likely that the different predictive power of the two measures of prior arrest history is caused by the significant correlation between the two variables. With a correlation coefficient of .74—a measure of quantitative association between two variables, ranging in values from 1 for perfect correlation to 0 for no relationship at all, the stability of the results could be affected by sample size and the presence of other collinear variables in the models.
7. These variables are also correlated with program termination but in the reverse direction.
8. David P. Farrington, “Predicting Individual Crime Rates” (1987) 9 *Crime and Justice* 53–101.
9. With offense types expanded to include all crimes against persons, the overall re-arrest rate increases to 23 percent. This does not suggest, however, that only 4 percent (23 minus 19) had committed an assault offense in which the victim is not specifically identified as either spouse or partner, as the re-arrest measures are created based on the first instance of re-arrest occurring after intake without accounting for possible re-arrests of other offense types during the subsequent follow-up period.
10. Edward Leamer and H. Leonard, “Reporting the Fragility of Regression Estimates” (1983) 65 *Review of Economics and Statistics* 306–317.





## Chapter 6: Policy Issues and Research Implications

### *Introduction*

Domestic violence represents both a serious criminal justice and public health problem. While the offenders in our sample are not representative of men who commit domestic violence in the population at large, they do possess many of the same characteristics that are typical of the population caught up in the criminal justice system more generally: low levels of educational attainment, marginal employment, minority status, prior criminal history, and a tendency for drug and alcohol abuse. The crimes committed by the men in the sample are unique, however, because they involve an intimate partner—someone with whom the offender often has an ongoing relationship that may include cohabitation, shared responsibilities for raising children, and/or co-mingled finances.

In part because of the special relationship between the offender and the victim in domestic violence cases, the criminal justice system has struggled to find an appropriate response. Current policy embodied in Pen. Code §1203.097 represents an effort by lawmakers to correct past failures of the justice system to recognize the severity of the problem of domestic violence and to hold offenders accountable. This policy represents a combination of both deterrence and rehabilitation: sanctions against offenders for failure to comply with terms of probation as deterrence and educational programs for rehabilitation. It remains unclear how effective either component of the policy is in achieving the ultimate goals of holding offenders accountable and increasing victim safety.

In this study, our conceptual framework incorporates multiple levels of variance, from the different characteristics of individual offenders to programs differences and system level variables, especially different forms of intervention by probation and the courts. Measurements of variance at each of these levels are based on different approaches that are largely determined by our ability to identify, define, and capture the relevant data.

We collected an extensive amount of data from multiple sources at the level of individual offenders, including both legal and extralegal variables. At the level of batterer intervention programs (BIPs), differences in program characteristics are measured by an “inventory” survey that describes various program philosophies, curriculum topics, and treatment approaches and practices. In terms of probation or court supervision of offenders, both the proportion and frequency under either form of supervision are measured across the jurisdictions.

When we created statistical models to examine the relative efficacy of different modes of intervention/monitoring strategies, for various reasons we were not able to incorporate all quantitative measures into the models. The limited set of factors included in the empirical statistical analyses necessarily leads to a simplified representation of reality. Jurisdictional comparisons are based on broad categories, without their constituent elements decomposed and connected to the relevant outcome measures. We treated differences in program characteristics as statistical variances without pointing to the qualitative dimension of their differences. At the

level of individual offenders, the study is invariably constrained, as in all quasi-experimental study designs, by observable and measurable variables.<sup>1</sup>

With the above limitations regarding data and measurement issues in mind, we can summarize the major findings of the study as follows:

- The men who find their way into the justice system and ultimately enroll in BIPs appear to be non-representative of the larger social problem of domestic violence. The sample of men convicted of domestic violence offenses drawn for this study generally had low levels of educational attainment, were poor, majority Hispanic, and had lengthy criminal records;
- Slightly more than one third of the men convicted of domestic violence in our sample report that they still live with their victim; about one third of the men reported that they live with children;
- BIPs appear to incorporate multiple approaches to intervention with domestic violence offenders into their programs, integrating components of cognitive behavioral therapy, the Duluth model and other methods that they determine are appropriate and effective;
- The educational topics that BIPs identified as important to helping offenders end their abuse appear to be consistent with the legislative requirements for these programs;
- Offenders' rates of program completion varied across different BIPs. The reason for this, however, appears to be in part that the characteristics of men who are enrolled in different BIPs varies systematically across programs. The statistical significance of the differences in program completion across BIPs declines as additional, individual-level variables are added to the model;
- In contrast to the weak correlation between program completion and BIP, there is no statistical association at all between programs and an offender's likelihood of re-offense;
- For offenders who successfully completed the 52-week BIP, attitudes and beliefs showed small, positive, changes along a number of dimensions including taking greater personal responsibility, understanding the effect of abuse on others, and anger management;
- The strongest predictors of whether or not men were re-arrested following intake in a BIP were individual characteristics of the offenders, not the characteristics of jurisdictions or BIPs in which offenders were enrolled.<sup>2</sup> Men who were more educated, older, had shorter criminal histories, and did not display clear signs of drug or alcohol dependence had a lower likelihood of re-arrest;
- Whether probation or the court is primarily responsible for oversight of the offenders made no difference in the likelihood of re-arrest. This finding is similar to the conclusion of a recent study in which judicial supervision of domestic violence offenders—with comparisons between supervision of different intensity and compared with no supervision

at all—was found not to make any significant difference on recidivism 12 months after sentencing;<sup>3</sup>

- Even after controlling for individual characteristics, two jurisdictions showed statistically significant differences in outcomes for offenders. Using Los Angeles as the base for comparison, offenders in Solano County had a likelihood of re-arrest at 12 months after intake that is one-third the likelihood of offenders in Los Angeles County, while offenders in Santa Clara County were 1.6 times as likely to be arrested as offenders in Los Angeles.

It bears repeating that the absence of statistically significant differences in the relative efficacy of probation versus court supervision, or among programs with different program philosophies and practices, needs to be understood in the context of the high-level and broad conceptual framework in which the jurisdictional and program features are defined, categorized, and measured. The “no difference” finding does not address the relative efficacy of any specific element that constitutes the “system,” nor the various programmatic elements that differentiate one program from another.

The similarities of outcomes across jurisdictions and the salience of individual variables in predicting outcomes may be caused by a number of factors related to the intervention itself or to the design of the research. As some of these factors may fall outside of the scope of the quantitative data collected and analyzed, this concluding section of the report draws upon the qualitative data gathered in the course of the research study from field observations and interviews—information that cannot easily be integrated into statistical models. Our concluding remarks are divided into two major sections, one that reflects issues of criminal justice policy and the other related to the research implications of our findings.

## ***Criminal Justice System Policy Issues***

### **Variation in Offender Characteristics May Allow for More Differentiated Case Management**

The pattern of findings emerging from this study suggests that there are common characteristics among justice system partners across and within court jurisdictions, as well as common social and psychological characteristics among batterers participating in the study. As noted above, the characteristics of many of the men in our sample suggest that the domestic violence cases that find their way into the justice system and end up in BIPs are multidimensional problems. Many of the offenders in the sample have problems with financial stability, attained low levels of education, have prior criminal histories, and struggle with issues of drug and alcohol abuse. Despite these similarities, our data also suggests that the following are important forms of variation across offenders:

- While male batterers generally appear to struggle with unemployment or underemployment, there are also statistically significant differences in the educational attainment and thus the literacy level of these men;

- There is a relationship between employment and educational status and domestic living situation, with male batterers residing in a number of different types of domestic arrangements;
- Prior criminal arrests are not uncommon among male batterers, but the age of onset and the length of this history appear to vary widely;
- While many male batterers may be at-risk for alcohol and drug abuse, there is variation in these scores as well.

Qualitative data gathered from departments of probation in our study jurisdictions indicate that in some of the jurisdictions only formally supervised offenders are assessed prior to assignment to a BIP. Even when misdemeanants were assessed, it was unclear that the information was systematically used for purposes of placing the offender in a specific BIP or for requiring treatment for drug and alcohol abuse or mental illness. Instead, intake protocols used by departments of probation, when they occur, appear focused primarily on *risk*-assessment rather than *needs*-assessment.

Given the importance of individual risk factors identified in this study, screening mechanisms should seek, to the extent possible, to include needs assessment to assist in directing offenders to resources that might improve their chances of successfully completing the BIP and remaining violence free during and following their attendance in the program.

## **Enhanced Intake/Assessment May Improve Offender Treatment**

Enhancing the needs assessment of offenders during intake would expand upon current law. Pen. Code §1203.097(b)(1) explicitly lays out a requirement for this type of assessment but limits it to offenders who are on formal probation. In these cases the probation department shall

make an investigation and take into consideration the defendant's age, medical history, employment and service records, educational background, community and family ties, prior incidents of violence, police report, treatment history, if any, demonstrable motivation, and other mitigating factors in determining which batterer's program would be appropriate for the defendant.

The penal code does not mandate a similar intake process for defendants who are sentenced to court-supervised or informal probation, and in three of the five study counties, the majority of offenders are informally supervised. However, in two locations, Riverside County and the Long Beach court in Los Angeles County, offenders under informal probation undergo a prescreening process with an intake component that is supervised by an outside party. In Riverside County, the non-profit Volunteer Center oversees the intake process, whereas in Long Beach the Public Health Office performs that function. Therefore, even at locations that do not supervise offenders formally, there may be mechanisms in place to ensure that every offender could be screened prior to enrollment in a BIP.

With clearer delineation of the risk factors associated with different offender populations, BIPs might be able to tailor their treatment more narrowly. The findings from the Program Content Survey (PCS), described more fully in Chapter 3, suggest that BIPs currently take a cross-disciplinary approach to their intervention with male batterers, with anecdotal reports from senior facilitators suggesting that this is necessary because a single treatment model simply does not capture the complex and varying needs, problems, and strengths of their clients and their partners. Further research, including consultation with BIP practitioners and those specializing in batterer intervention would be needed to develop this concept further.

## **Drug/Alcohol Treatment May Be Important to Help Offenders End Their Abuse**

Many male batterers participating in this study indicate through their CAGE scores that they are at risk for alcohol and drug abuse, with anecdotal information from interviews with program staff suggesting that the incidence of this problem is even higher than the CAGE scores reveal. Further, higher CAGE scores are robust predictors of non-completion of batterer intervention programs, and senior program staff responding to the PCS point out that addressing the topic of alcohol and drug abuse is important in helping their clients end their domestic abuse.

Given the current mandate for domestic violence treatment outlined in Pen. Code §1203.097, the limited resources available to most BIPs in California as well as the limited leverage that they may exercise over offenders, it may be useful for departments of probation and the courts to consider how best to support BIPs in requiring batterers at risk for substance abuse to attend some reasonable form of drug/alcohol treatment in conjunction with their enrollment in the BIP.

## **Current BIP Fee Structure May Hinder Differentiated Case Management**

One more piece of the puzzle of differentiated case management has to do with fees. The fees paid by batterers are designed to hold offenders accountable for their domestic violence, promote their sense of investment in the programs in which they enroll, and sustain intervention programs financially for their intervention work with enrollees. Generally paid on a sliding scale by batterers, the fees often represent only partial compensation for the costs of the intervention program. Nonpayment of fees was frequently cited as a reason for program termination and/or failure to complete the BIP, and the collection of fees sometimes appears to absorb a significant amount of the program staff's time and effort.

The current method of assessing and paying fees, all managed at the BIP level, may pose a barrier to a differentiated treatment model because Pen. Code §1203.097 mandates probation departments to evenly allocate referrals of indigent clients among approved programs. Thus, the effort to assign the right socioeconomic balance to different programs might very well undermine efforts to assign men to programs on the basis of the characteristics that put them most at risk for re-offense.

Moreover, given the predominance of lower-income men in these groups, it is not clear that enough differentiation exists along the dimension of income to sustain BIPs. More often than not, the BIP appears to make the final evaluation as to whether or not it can absorb another

indigent client into the program. Creating a more differentiated treatment model might require an exploration of alternative fee distribution and payment plans. This might grant BIPs the financial freedom to accept enrollments on the basis of service need rather than have to consider a client's ability to pay.

## ***Research Implications***

### **Systems Analysis Hampered by Variation *within* Jurisdictions**

The first and perhaps most challenging of the findings as they relate to the methodology and implications for future research is the simple fact that the systems analysis that we sought to conduct was frequently undermined by the lack of “systemness” within jurisdictions. Differences in court practice from location to location within jurisdictions, as well as large variability in outcomes across BIPs within jurisdictions, undercut our efforts to evaluate the justice system response. Instead, in some cases we have findings related to different systems within a single jurisdiction.

Further integration of the qualitative data will assist with the interpretation of the findings. Once the qualitative differences within jurisdictions are better understood, quantitative analysis that excludes outlying court locations where these introduce too much variability might be a fruitful path for recapturing the system perspective that motivated this study. Given the clustering of large numbers of offenders in specific courts and in some specific BIPs, this may be a near- to medium-term follow-up with this data set.

Also, system intervention, measured as “probation contact,” “court review,” or even “attendance” at the BIP are all limited measures. Consistent with the other observations here, more qualitative information on what these variables really are in practice—whether probation contact is a face-to-face interview at the department of probation as opposed to a check-in by telephone or whether the review at the trial court is in open court in front of a judge or handled by a courtroom clerk—would assist in distinguishing among different systems.

### **More Information on BIPs Is Needed to Understand and Identify Promising Practices**

In addition to the challenge presented by variability within individual jurisdictions, to some extent the BIPs remain black boxes. While the PCS captured valuable information related to the priorities for teaching and training that program facilitators attach to different elements of the intervention, it did not identify sufficient variability to introduce the data into our quantitative models and to begin teasing out the effects that these programs produce on offender outcomes.

In the future, this information will need to be triangulated with independent forms of data if we are to clearly understand the approach intervention programs are taking in their work with clients. Further, we need to learn more about BIPs as practitioner groups and/or organizations in terms of their staffing levels and role differentiation, the training and professional experience levels of program staff, the supplementary services BIPs are able to provide clients directly or indirectly, and the resources these organizations have at their disposal to sustain their work with

batterers. Such information is essential to our ability to open up the black box of the BIPs in their various organizational forms, as well as to identify promising program approaches and practices.

### **More Refined Psychosocial Measures of Individuals Are Needed**

Additional data at the level of individual offenders may also be needed. While our measures of socioeconomic characteristics and criminal histories appear to differentiate offenders in the sample sufficiently to control for these factors, our data on the psychosocial characteristics of individuals is less robust. The Conflict Tactics Scale 2 (CTS2), in particular, did not detect differences among offenders, leaving considerable uncertainty about individual offenders in terms of their history of abusive behavior. Constraints on the CTS2 included the time period that BIPs could provide us for the assessment of each batterer, the particular session in which this assessment must occur, and the form it must take. Moreover, the self-reporting nature of this instrument coupled with its administration in a time-constrained intake session led to what appears to be various forms of suppression effect and response bias in the answers of new program enrollees.

This combination of factors severely limited the usefulness of the information, and along with other findings of this study suggests the need for a more in-depth assessment of batterers at a time and in a setting where they may provide more accurate responses. Further, any form of assessment that is undertaken should probably allow for a deeper understanding of the psychosocial profile of batterers, as well as sufficient contextual information about their life situations, to allow both practitioners and researchers to better understand them with the ultimate goal of preventing future domestic violence.

And, while the BIP Process Survey provided a useful tool for examining the impact of batterer intervention programs on the offender's attitudes and beliefs, instead of being limited to program completion and re-offense as outcome indicators, additional research that further refines and tests the instrument in different study contexts would improve our understanding of the psychosocial aspects of domestic violence behavior and the intervention programs.

Further analysis will also need to be conducted on the causal connection between psychosocial changes observed in the BIP Process Survey and behavioral changes as they relate to domestic violence. Given the lengthy criminal histories that many men in the study have as well as other risk factors that we identified, the finding of positive, statistically significant changes in attitudes and beliefs—however slight—should not be discounted. A more comprehensive theoretical understanding of these indicators, though, will be useful for future analysis.

### ***Other Issues***

At this point, the follow-up period for the study is necessarily short due to the time frame of the grant, but the data should be revisited and examined again at a later date for longer follow-up. An observation made by a number of the most senior clinicians participating in this study was that if the batterer had a significant history of domestic violence either as a childhood victim or an adult perpetrator, one should anticipate that change in abusive behavior would take time. More specifically, while we might anticipate that attitudes, beliefs, and behavior may begin to shift in a more pro-social direction in the first year of program assignment, deeper forms of



change in these domains may take years, often in conjunction with periodic contact and even reenrollment in programs aimed at reforming offenders.

This suggests the need to continue to follow graduates of these programs over a longer period of time than is permitted by the resources available in this study, particularly if we are to understand the full trajectory of change in these individuals, as well as the support services that may be necessary to sustain this change.

Finally, in our examination of patterns of re-arrests as an outcome indicator, the length of follow-up period is closely related to the time during follow-up when the offender is at large in the community versus being incapacitated in some form (detention, jail, treatment institution, or prison). The status of the offender and the amount of time in this status, thus, affect the offender's opportunity for re-offense and re-arrest.<sup>4</sup> Thus, distinguishing between an incapacitation effect and a treatment effect is not possible with the current data set.

For an offender incarcerated for a substantial period of time during follow-up, the lack of re-offense reflects the direct impact of incapacitation effect, rather than any treatment effect from the batterer intervention program. The present study did not track the offenders' incarceration records as part of the follow-up analysis. As the likelihood for incapacitation may vary across the jurisdictions, reflecting partly the different sanctions applied by the courts and probation as well as the different overcrowding situation in local jails, future studies need to control for the potential confounding effect from incapacitation in order to better understand the system impact.

## Endnotes: Chapter 6

---

<sup>1</sup> Edward E. Leamer, “Let’s Take the Con Out of Econometrics” (March 1983) 73(1) *The American Economic Review* 31–43.

<sup>2</sup> All findings discussed in this Executive Summary are statistically significant at a level of .01 or .05 unless otherwise noted.

<sup>3</sup> Melissa Labriola, Michael Rempel, and Robert C. Davis, *Testing the Effectiveness of Batterer Programs and Judicial Monitoring*, Center for Court Innovation (November 2005).

<sup>4</sup> Daniel F. McCaffrey, Andrew R. Morral, Greg Ridgeway, and Beth Ann Griffin, “Interpreting Treatment Effects When Cases Are Institutionalized After Treatment” (2007) 89 *Drug and Alcohol Dependence* 126–138.



## Appendix A: Partial list of interviews conducted

| Date               | Institution                                              | County                |
|--------------------|----------------------------------------------------------|-----------------------|
| November 17, 2005  | Court / Information Systems                              | Los Angeles           |
| November 18, 2005  | Court / Calendar                                         | Los Angeles           |
| November 18, 2005  | BIPs / Presentation & Interview                          | Los Angeles           |
| December 20, 2005  | BIP                                                      | Los Angeles           |
| January 26, 2006   | Probation                                                | Riverside             |
| February 7, 2006   | Probation                                                | Riverside             |
| February 7, 2006   | "Volunteer Center"                                       | Riverside             |
| February 28, 2006  | County Government / Information Systems                  | San Joaquin           |
| February 28, 2006  | Probation & BIPs                                         | San Joaquin           |
| April 27, 2006     | BIP / Group Observation                                  | San Joaquin           |
| September 6, 2006  | Probation                                                | Riverside             |
| September 25, 2006 | Probation                                                | Santa Clara           |
| July 21, 2006      | BIP                                                      | Santa Clara           |
| August 1, 2006     | BIP                                                      | Santa Clara           |
| August 22, 2006    | BIP                                                      | San Joaquin           |
| November 20, 2006  | BIP                                                      | San Joaquin/Solano    |
| December 18, 2006  | BIP                                                      | San Joaquin           |
| February 5, 2007   | Probation                                                | Santa Clara           |
| March 17, 2007     | Probation                                                | Los Angeles           |
| March 20, 2007     | Probation                                                | Riverside             |
| April 24, 2007     | Probation                                                | Los Angeles           |
| June 18, 2007      | BIP                                                      | Santa Clara           |
| June 22, 2007      | Probation                                                | Los Angeles           |
| June 26, 2007      | Court                                                    | Solano                |
| June 26, 2007      | Probation                                                | Solano                |
| July 9, 2007       | Court                                                    | San Joaquin           |
| July 9, 2007       | Probation                                                | San Joaquin           |
| August 9, 2007     | Court                                                    | Riverside             |
| August 23, 2007    | Court                                                    | Santa Clara           |
| October 19, 2007   | Court / Probation / BIPs                                 | Riverside             |
| October 31, 2007   | Court                                                    | San Joaquin           |
| November 29, 2007  | BIP                                                      | San Joaquin           |
| February 27, 2008  | Court                                                    | Solano                |
| March 7, 2008      | Court                                                    | Los Angeles           |
| March 10, 2008     | Court                                                    | Los Angeles           |
| May 7, 2008        | Court / Probation / BIPs / Other Justice System Agencies | Riverside / Temecula  |
| May 30, 2008       | Court / Probation / BIPs / Other Justice System Agencies | Riverside / Riverside |
| June 4, 2008       | Court / Probation                                        | Los Angeles           |
| June 13, 2008      | Court / Probation / BIPs / Other Justice System Agencies | Santa Clara           |
| June 16, 2008      | Probation / BIPs                                         | San Joaquin           |
| August 13, 2008    | Court / Probation                                        | Solano                |

## Appendix B: Supplemental Information Form

|                                           |
|-------------------------------------------|
| Date:                                     |
| Name:                                     |
| Probation Case #:                         |
| Court Case #:                             |
| CTS2 Intake Form #<br>(at bottom of CTS2) |

### Client's Supplemental Information Form (For Completion by Program Staff)

**Directions to Staff Person:** Please fill out this supplemental information sheet as completely as possible. If you need to refer to other forms of intake information please feel free to do so. Remember to staple this information sheet to the CTS2 once the client has finished filling out his form. Please mail all of this information back to the Office of Court Research in the stamped, self addressed envelopes provided.

**Please check the box or fill in the answer that comes closest to describing your client:**

1. Education level (Highest level completed)

- ☐ Less than high school
- ☐ High School Diploma / GED
- ☐ Some College / Tech School / A.A. Degree
- ☐ College Graduate
- ☐ Graduate or Professional Degree

2. Primary ethnicity

- ☐ Asian or Pacific Islander
- ☐ African American or Black
- ☐ American Indian or Alaskan Native
- ☐ Hispanic or Latino
- ☐ White

☐ Other, Please Describe: \_\_\_\_\_

3. What did this client indicate his income was for the previous year? (If he indicated an income range please list the range) \_\_\_\_\_

4. What is the relationship of the victim to the client?

- ☐ Former wife
- ☐ Former girlfriend
- ☐ Current wife
- ☐ Current girlfriend

5. Are the client and this person currently living together?

- ☐ No
- ☐ Yes

## Appendix B: Supplemental Information Form

6. Does the client currently have children of his own that are living with him, living nearby, or living some distance away that he sees regularly?

- ☐ Client does not have children.
- ☐ Client has children that live with him.
- ☐ Client has children that do not live with him, but he visits them once a month or more.
- ☐ Client has children, but does not visit them regularly (less than once a month).

7. Has this client received any of the following services during the following time periods? Please check the appropriate box for each type of service.

|                                 | Currently<br>enrolled    | Yes, in the<br>previous 12<br>months | Yes, more<br>than a year<br>ago | Never                    | Don't<br>Know            |
|---------------------------------|--------------------------|--------------------------------------|---------------------------------|--------------------------|--------------------------|
| Alcohol or drug treatment       | <input type="checkbox"/> | <input type="checkbox"/>             | <input type="checkbox"/>        | <input type="checkbox"/> | <input type="checkbox"/> |
| Anger Management                | <input type="checkbox"/> | <input type="checkbox"/>             | <input type="checkbox"/>        | <input type="checkbox"/> | <input type="checkbox"/> |
| Counseling or therapy           | <input type="checkbox"/> | <input type="checkbox"/>             | <input type="checkbox"/>        | <input type="checkbox"/> | <input type="checkbox"/> |
| Batterer's Intervention Program | <input type="checkbox"/> | <input type="checkbox"/>             | <input type="checkbox"/>        | <input type="checkbox"/> | <input type="checkbox"/> |
| Parenting class or training     | <input type="checkbox"/> | <input type="checkbox"/>             | <input type="checkbox"/>        | <input type="checkbox"/> | <input type="checkbox"/> |

8. If the client has been enrolled in a BIP in the last year, for how many weeks did he attend? (Please add all of the course sessions together regardless of the number of courses.)

---

9. What *weekly* enrollment fee is this client currently paying? \_\_\_\_\_

10. How accurate do you think the information is that this client provided?

- ☐ Accurate
- ☐ Somewhat inaccurate
- ☐ Highly inaccurate
- ☐ Very difficult to say

11. If the information seems inaccurate, please explain:

|  |
|--|
|  |
|--|

## Appendix C: Revised Conflict Tactics Scale 2 and CAGE Assessment

### Office Use Only

Date: \_\_\_\_\_

Probation Case #: \_\_\_\_\_

Court Case #: \_\_\_\_\_

Name: \_\_\_\_\_

### CTS2 Behavior in Relationships

**Introduction:** In the following survey you will be asked some questions about what may have happened when you and your partner had disagreements or disputes about things in the last year. These questions will be about your behavior during these disputes, although we know that this may not represent a complete picture of what happened. On the other hand your answers to these questions will help us understand how you have handled these disagreements in the past, and how the program may help you find new ways of dealing with disagreements with your partner in the future.

Please think of how you have dealt with your partner over the last 12 months as you answer the following questions. While we want you to answer each accurately and truthfully, don't think too much about any single question. Just give us your best estimate of how often things have happened and move on to the next question.

If one of these things did not happen in the past year, but it happened before then go ahead and circle the number "7".

|                                                                                                            | Once | Twice | 3-5<br>times | 6 - 10<br>times | 11 - 20<br>times | More<br>than<br>20<br>times | Not in<br>the<br>past<br>year | Never |
|------------------------------------------------------------------------------------------------------------|------|-------|--------------|-----------------|------------------|-----------------------------|-------------------------------|-------|
| 1. I showed my partner I cared even though we disagreed.                                                   | 1    | 2     | 3            | 4               | 5                | 6                           | 7                             | 0     |
| 2. I explained my side of a disagreement to my partner.                                                    | 1    | 2     | 3            | 4               | 5                | 6                           | 7                             | 0     |
| 3. I insulted or swore at my partner.                                                                      | 1    | 2     | 3            | 4               | 5                | 6                           | 7                             | 0     |
| 4. I threw something at my partner that could hurt.                                                        | 1    | 2     | 3            | 4               | 5                | 6                           | 7                             | 0     |
| 5. I twisted my partner's arm or hair.                                                                     | 1    | 2     | 3            | 4               | 5                | 6                           | 7                             | 0     |
| 6. I had a sprain, bruise, or small cut because of a fight with my partner.                                | 1    | 2     | 3            | 4               | 5                | 6                           | 7                             | 0     |
| 7. I showed respect for my partner's feelings about an issue.                                              | 1    | 2     | 3            | 4               | 5                | 6                           | 7                             | 0     |
| 8. I made my partner have sex without a condom.                                                            | 1    | 2     | 3            | 4               | 5                | 6                           | 7                             | 0     |
| 9. I pushed or shoved my partner.                                                                          | 1    | 2     | 3            | 4               | 5                | 6                           | 7                             | 0     |
| 10. I used force (like hitting, holding down, or using a weapon) to make my partner have oral or anal sex. | 1    | 2     | 3            | 4               | 5                | 6                           | 7                             | 0     |
| 11. I used a knife or gun on my partner.                                                                   | 1    | 2     | 3            | 4               | 5                | 6                           | 7                             | 0     |

## Appendix C: Revised Conflict Tactics Scale 2 and CAGE Assessment

|                                                                                               | Once | Twice | 3-5 times | 6 - 10 times | 11 - 20 times | More than 20 times | Not in the past year | Never |
|-----------------------------------------------------------------------------------------------|------|-------|-----------|--------------|---------------|--------------------|----------------------|-------|
| 12. I passed out from being hit on the head by my partner in a fight.                         | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 13. I called my partner fat or ugly.                                                          | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 14. I punched or hit my partner with something that could hurt.                               | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 15. I destroyed something belonging to my partner.                                            | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 16. I went to a doctor because of a fight with my partner.                                    | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 17. I choked my partner.                                                                      | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 18. I shouted or yelled at my partner.                                                        | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 19. I slammed my partner against a wall.                                                      | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 20. I said I was sure we could work out a problem.                                            | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 21. I needed to see a doctor because of a fight with my partner, but I didn't.                | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 22. I beat up my partner.                                                                     | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 23. I grabbed my partner.                                                                     | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 24. I used force (like hitting, holding down, or using a weapon. to make my partner have sex. | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 25. I stomped out of the room or house or yard during a disagreement.                         | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 26. I insisted on sex when my partner did not want to (But did not use physical force).       | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 27. I slapped my partner.                                                                     | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 28. I had a broken bone from a fight with my partner.                                         | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 29. I used threats to make my partner have oral or anal sex.                                  | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 30. I suggested a compromise to a disagreement.                                               | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 31. I burned or scalded my partner on purpose.                                                | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 32. I insisted my partner have oral or anal sex (but did not use physical force).             | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |



## Appendix C: Revised Conflict Tactics Scale 2 and CAGE Assessment

|                                                                                           | Once | Twice | 3-5 times | 6 - 10 times | 11 - 20 times | More than 20 times | Not in the past year | Never |
|-------------------------------------------------------------------------------------------|------|-------|-----------|--------------|---------------|--------------------|----------------------|-------|
| 33. I accused my partner of being a lousy lover.                                          | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 34. I did something to spite my partner.                                                  | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 35. I threatened to hit or throw something at my partner.                                 | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 36. I felt physical pain that still hurt the next day because of a fight with my partner. | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 37. I kicked my partner.                                                                  | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 38. I used threats to make my partner have sex.                                           | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |
| 39. I agreed to try a solution to a disagreement my partner suggested.                    | 1    | 2     | 3         | 4            | 5             | 6                  | 7                    | 0     |

Please circle the number that comes closest to describing things in the last year.

| In the last year:                                                                                                     | Yes | No |
|-----------------------------------------------------------------------------------------------------------------------|-----|----|
| 40. Have you felt you should cut down on your drinking or drug use?                                                   | 1   | 2  |
| 41. Have people annoyed you by criticizing your drinking or drug use?                                                 | 1   | 2  |
| 42. Have you felt bad or guilty about your drinking or drug use?                                                      | 1   | 2  |
| 43. Have you had a drink or taken drugs first thing in the morning to steady your nerves or to get rid of a hangover? | 1   | 2  |
| 44. Have you lost your job or had your hours at work greatly reduced?                                                 | 1   | 2  |

45. What is your *current* employment status? (Please circle the alternative that is closest.)

- Employed full time for pay. 1
- Employed part-time for pay. 2
- Not employed for pay. 3

## Appendix D: Attendance Log

Program Name: \_\_\_\_\_

Client Name: \_\_\_\_\_ Probation Case #: \_\_\_\_\_

Intake Date: \_\_\_\_\_ Court Case #: \_\_\_\_\_

### ATTENDANCE LOG

**Instructions:** Record the client's attendance by marking each date the client was scheduled to attend, using the following key:

- ☐ Client attended the session
- ☒ Client did not attend- absence was excused
- ☒ Client did not attend- absence was **not** excused
- ☒ Client was terminated from program
- ☐ Date client reinstated into program, if applicable (mark subsequent dates of attendance or absence with circles and x's as shown above).

#### January 2006

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
| 1  | 2  | 3  | 4  | 5  | 6  | 7  |
| 8  | 9  | 10 | 11 | 12 | 13 | 14 |
| 15 | 16 | 17 | 18 | 19 | 20 | 21 |
| 22 | 23 | 24 | 25 | 26 | 27 | 28 |
| 29 | 30 | 31 |    |    |    |    |

#### February 2006

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    |    |    | 1  | 2  | 3  | 4  |
| 5  | 6  | 7  | 8  | 9  | 10 | 11 |
| 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 19 | 20 | 21 | 22 | 23 | 24 | 25 |
| 26 | 27 | 28 |    |    |    |    |

#### March 2006

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    |    |    | 1  | 2  | 3  | 4  |
| 5  | 6  | 7  | 8  | 9  | 10 | 11 |
| 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 19 | 20 | 21 | 22 | 23 | 24 | 25 |
| 26 | 27 | 28 | 29 | 30 | 31 |    |

#### April 2006

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    |    |    |    |    |    | 1  |
| 2  | 3  | 4  | 5  | 6  | 7  | 8  |
| 9  | 10 | 11 | 12 | 13 | 14 | 15 |
| 16 | 17 | 18 | 19 | 20 | 21 | 22 |
| 23 | 24 | 25 | 26 | 27 | 28 | 29 |
| 30 |    |    |    |    |    |    |

#### May 2006

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
| 1  | 2  | 3  | 4  | 5  | 6  |    |
| 7  | 8  | 9  | 10 | 11 | 12 | 13 |
| 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 21 | 22 | 23 | 24 | 25 | 26 | 27 |
| 28 | 29 | 30 | 31 |    |    |    |

#### June 2006

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    |    |    |    | 1  | 2  | 3  |
| 4  | 5  | 6  | 7  | 8  | 9  | 10 |
| 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| 25 | 26 | 27 | 28 | 29 | 30 |    |

#### July 2006

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    |    |    |    |    |    | 1  |
| 2  | 3  | 4  | 5  | 6  | 7  | 8  |
| 9  | 10 | 11 | 12 | 13 | 14 | 15 |
| 16 | 17 | 18 | 19 | 20 | 21 | 22 |
| 23 | 24 | 25 | 26 | 27 | 28 | 29 |
| 30 | 31 |    |    |    |    |    |

#### August 2006

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    |    | 1  | 2  | 3  | 4  | 5  |
| 6  | 7  | 8  | 9  | 10 | 11 | 12 |
| 13 | 14 | 15 | 16 | 17 | 18 | 19 |
| 20 | 21 | 22 | 23 | 24 | 25 | 26 |
| 27 | 28 | 29 | 30 | 31 |    |    |

#### September 2006

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    |    |    |    |    | 1  | 2  |
| 3  | 4  | 5  | 6  | 7  | 8  | 9  |
| 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| 17 | 18 | 19 | 20 | 21 | 22 | 23 |
| 24 | 25 | 26 | 27 | 28 | 29 | 30 |

## Appendix D: Attendance Log

### October 2006

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
| 1  | 2  | 3  | 4  | 5  | 6  | 7  |
| 8  | 9  | 10 | 11 | 12 | 13 | 14 |
| 15 | 16 | 17 | 18 | 19 | 20 | 21 |
| 22 | 23 | 24 | 25 | 26 | 27 | 28 |
| 29 | 30 | 31 |    |    |    |    |

### November 2006

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    |    |    | 1  | 2  | 3  | 4  |
| 5  | 6  | 7  | 8  | 9  | 10 | 11 |
| 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 19 | 20 | 21 | 22 | 23 | 24 | 25 |
| 26 | 27 | 28 | 29 | 30 |    |    |

### December 2006

| Su | Mo | Tu | We | Th | Fr | Sa  |
|----|----|----|----|----|----|-----|
|    |    |    |    |    |    | 1 2 |
| 3  | 4  | 5  | 6  | 7  | 8  | 9   |
| 10 | 11 | 12 | 13 | 14 | 15 | 16  |
| 17 | 18 | 19 | 20 | 21 | 22 | 23  |
| 24 | 25 | 26 | 27 | 28 | 29 | 30  |
| 31 |    |    |    |    |    |     |

### January 2007

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    | 1  | 2  | 3  | 4  | 5  | 6  |
| 7  | 8  | 9  | 10 | 11 | 12 | 13 |
| 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 21 | 22 | 23 | 24 | 25 | 26 | 27 |
| 28 | 29 | 30 | 31 |    |    |    |

### February 2007

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    |    |    |    | 1  | 2  | 3  |
| 4  | 5  | 6  | 7  | 8  | 9  | 10 |
| 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| 25 | 26 | 27 | 28 |    |    |    |

### March 2007

| Su | Mo | Tu | We | Th | Fr | Sa  |
|----|----|----|----|----|----|-----|
|    |    |    |    |    | 1  | 2 3 |
| 4  | 5  | 6  | 7  | 8  | 9  | 10  |
| 11 | 12 | 13 | 14 | 15 | 16 | 17  |
| 18 | 19 | 20 | 21 | 22 | 23 | 24  |
| 25 | 26 | 27 | 28 | 29 | 30 | 31  |

### April 2007

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
| 1  | 2  | 3  | 4  | 5  | 6  | 7  |
| 8  | 9  | 10 | 11 | 12 | 13 | 14 |
| 15 | 16 | 17 | 18 | 19 | 20 | 21 |
| 22 | 23 | 24 | 25 | 26 | 27 | 28 |
| 29 | 30 |    |    |    |    |    |

### May 2007

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    |    | 1  | 2  | 3  | 4  | 5  |
| 6  | 7  | 8  | 9  | 10 | 11 | 12 |
| 13 | 14 | 15 | 16 | 17 | 18 | 19 |
| 20 | 21 | 22 | 23 | 24 | 25 | 26 |
| 27 | 28 | 29 | 30 | 31 |    |    |

### June 2007

| Su | Mo | Tu | We | Th | Fr | Sa  |
|----|----|----|----|----|----|-----|
|    |    |    |    |    |    | 1 2 |
| 3  | 4  | 5  | 6  | 7  | 8  | 9   |
| 10 | 11 | 12 | 13 | 14 | 15 | 16  |
| 17 | 18 | 19 | 20 | 21 | 22 | 23  |
| 24 | 25 | 26 | 27 | 28 | 29 | 30  |

### July 2007

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
| 1  | 2  | 3  | 4  | 5  | 6  | 7  |
| 8  | 9  | 10 | 11 | 12 | 13 | 14 |
| 15 | 16 | 17 | 18 | 19 | 20 | 21 |
| 22 | 23 | 24 | 25 | 26 | 27 | 28 |
| 29 | 30 | 31 |    |    |    |    |

### August 2007

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    |    |    | 1  | 2  | 3  | 4  |
| 5  | 6  | 7  | 8  | 9  | 10 | 11 |
| 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 19 | 20 | 21 | 22 | 23 | 24 | 25 |
| 26 | 27 | 28 | 29 | 30 | 31 |    |

### September 2007

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    |    |    |    |    |    | 1  |
| 2  | 3  | 4  | 5  | 6  | 7  | 8  |
| 9  | 10 | 11 | 12 | 13 | 14 | 15 |
| 16 | 17 | 18 | 19 | 20 | 21 | 22 |
| 23 | 24 | 25 | 26 | 27 | 28 | 29 |
| 30 |    |    |    |    |    |    |

### October 2007

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    | 1  | 2  | 3  | 4  | 5  | 6  |
| 7  | 8  | 9  | 10 | 11 | 12 | 13 |
| 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 21 | 22 | 23 | 24 | 25 | 26 | 27 |
| 28 | 29 | 30 | 31 |    |    |    |

### November 2007

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    |    |    |    | 1  | 2  | 3  |
| 4  | 5  | 6  | 7  | 8  | 9  | 10 |
| 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| 25 | 26 | 27 | 28 | 29 | 30 |    |

### December 2007

| Su | Mo | Tu | We | Th | Fr | Sa |
|----|----|----|----|----|----|----|
|    |    |    |    |    |    | 1  |
| 2  | 3  | 4  | 5  | 6  | 7  | 8  |
| 9  | 10 | 11 | 12 | 13 | 14 | 15 |
| 16 | 17 | 18 | 19 | 20 | 21 | 22 |
| 23 | 24 | 25 | 26 | 27 | 28 | 29 |
| 30 | 31 |    |    |    |    |    |

## Appendix E: BIP Process Survey

### Instructions for the BIP Process Survey

**Background.** The BIP Process Survey is designed to assess psychosocial change in a client as a result of enrollment in a batterer intervention program. Our objective is to learn more about how clients participating in the current study think about their relationship with their partners, as well as how they experience any interpersonal conflict that may occur. Toward this end your clients will be asked to indicate how much they agree or disagree with a series of statements in the survey.

Since this instrument is designed to track program impact on a psychosocial level, it should be administered twice for each client participating in the study. The first administration should occur toward the end of the first month of client enrollment, while the second should occur at the end of the 52 week intervention program. Please have client name and other ID information filled out on each survey for clients participating in the OCR study.

**Administration procedure.** Please administer the BIP Process Survey to a client after he has attended 4 to 5 weeks of class in your program. The BIP Process survey will take about 10 to 15 minutes to complete.

It is important to administer these surveys at a time that promotes accurate and complete responses, while also taking into consideration what is convenient for you and your clients. For example, if you think it best to administer this survey individually or in groups please do so. In all cases it is important to insure that clients will be undistracted by other things while completing the survey and that a staff person will be available to answer questions that may arise. Please remind clients that they should circle the number corresponding to their level of agreement with each statement in the survey.

It is also important for you to inform your clients that:

- There are no correct or incorrect responses to survey items. A client should simply respond to a question with his best sense of what describes his current views and experiences.
- All client responses are confidential. None of his individual responses will be shared with the courts, probation, or any other agency or group, nor will he be personally identifiable.

**Survey administration on an individual basis.** A number of methods may be considered when administering the BIP Process Survey to individual clients. For example, you may want to have clients fill out the survey:

- Before a regularly scheduled session by arriving early
- After a regularly scheduled group meeting or individual session
- During a regularly scheduled group meeting by briefly pulling a client from a group session

**Survey administration in groups.** If it is necessary to administer the survey in a group please make sure that clients participating in the study do not revise their answers as a result of these conversations.

**Analysis of BIP Process Survey data.** For important methodological reasons only those surveys corresponding to clients participating in the present OCR study may be processed and analyzed by the OCR. However, please forward all completed surveys to the OCR.

**Clients' names and ID numbers.** It is important to assign the proper client name and Court or Probation ID number to surveys completed by a clients participating in the OCR study. This will allow us to merge responses to the BIP Process Survey with client data that we have previously gathered. You may refer to the client rosters provided with the CTS2 to identify clients participating in the OCR study.

**Forwarding completed surveys to the OCR.** Please forward all surveys completed by clients to the OCR in the envelopes provided. You should forward these surveys at the beginning of each month.

## **Appendix E: BIP Process Survey**

**Copyright and limits of use:** Please do not use the BIP Process Survey for any purpose other than the present OCR study.

## Appendix E: BIP Process Survey

### BIP Process Survey

Please take a few moments to think about specific violent or abusive conflicts you have had with your partner. Now, based on these memories, please indicate the extent to which you agree or disagree with each of the following statements.

|                                                                                            | Strongly<br>Disagree | Disagree | Slightly<br>Disagree | Slightly<br>Agree | Agree | Strongly<br>Agree |
|--------------------------------------------------------------------------------------------|----------------------|----------|----------------------|-------------------|-------|-------------------|
| 1. I have control over whether I am abusive.                                               | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 2. I am responsible for my abusive behavior.                                               | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 3. If I'm upset, I usually take it out on my partner.                                      | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 4. In a conflict with my partner, I usually get what I want.                               | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 5. My abusive behavior has caused my family members to trust me less.                      | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 6. I am dependent on my partner.                                                           | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 7. My partner's behavior forces me to act abusively.                                       | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 8. I feel powerless during conflicts with my partner.                                      | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 9. When I am abusive, I feel that I am not under control of myself.                        | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 10. Taking a break helps me manage my anger.                                               | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 11. People in my life have been strongly affected by my abusive behavior.                  | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 12. I worry that my partner is going to leave me.                                          | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 13. I am in control of how I respond to my partner.                                        | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 14. I have lost relationships due to my abusive behavior.                                  | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 15. I can control my anger during conflicts with my partner.                               | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 16. When I don't have the final say in discussions with my partner, I feel out of control. | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 17. My abusive behavior has had long lasting effects on my family members.                 | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 18. I don't know what I would do without my partner.                                       | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 19. When I feel good about myself, I'm less likely to get into arguments.                  | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 20. I can express my anger without becoming abusive.                                       | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 21. Thinking positively about myself helps me avoid becoming abusive.                      | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 22. I worry about losing my relationship with my partner.                                  | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 23. I am responsible for the effects my abusive behavior has on others.                    | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 24. When I am becoming angry, I can feel it in my body.                                    | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 25. The only person I can control is me.                                                   | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 26. When my partner disagrees with me, I feel alone.                                       | 1                    | 2        | 3                    | 4                 | 5     | 6                 |
| 27. I'm responsible for my own happiness.                                                  | 1                    | 2        | 3                    | 4                 | 5     | 6                 |

## Appendix E: BIP Process Survey

|                                                                                      |   |   |   |   |   |   |
|--------------------------------------------------------------------------------------|---|---|---|---|---|---|
| 28. My abusive behavior has caused my family members to feel bad about themselves.   | 1 | 2 | 3 | 4 | 5 | 6 |
| 29. I feel jealous when my partner spends too much time with other people.           | 1 | 2 | 3 | 4 | 5 | 6 |
| 30. I have a choice about whether I am abusive or not.                               | 1 | 2 | 3 | 4 | 5 | 6 |
| 31. My abusive behavior has hurt me.                                                 | 1 | 2 | 3 | 4 | 5 | 6 |
| 32. I use violence to help me get what I want from my partner.                       | 1 | 2 | 3 | 4 | 5 | 6 |
| 33. When my partner does something without me, I feel left out.                      | 1 | 2 | 3 | 4 | 5 | 6 |
| 34. I feel better about my relationship with my partner when I'm the one in control. | 1 | 2 | 3 | 4 | 5 | 6 |
| 35. I know when I'm about to explode.                                                | 1 | 2 | 3 | 4 | 5 | 6 |
| 36. The main reason I'm in this group is because I have to be .                      | 1 | 2 | 3 | 4 | 5 | 6 |
| 37. My happiness typically depends on my partner.                                    | 1 | 2 | 3 | 4 | 5 | 6 |
| 38. When I have a bad day, I take it out on people at home.                          | 1 | 2 | 3 | 4 | 5 | 6 |
| 39. I am not responsible for my actions when I get in a rage.                        | 1 | 2 | 3 | 4 | 5 | 6 |
| 40. I need my partner to make me happy.                                              | 1 | 2 | 3 | 4 | 5 | 6 |
| 41. I know when I'm getting angry.                                                   | 1 | 2 | 3 | 4 | 5 | 6 |
| 42. I would come to this program even if I was not required to.                      | 1 | 2 | 3 | 4 | 5 | 6 |

### Duluth Model

The Domestic Abuse Intervention Project of Duluth Minnesota has given rise to a model of domestic violence intervention that has proven highly influential in California over the last twenty-five years. Commonly referred to as the Duluth Model, it calls for a justice system intervention into this syndrome that is designed to orchestrate responses by a community's point of first emergency contact with a community's women's shelters, police department, district attorney's office, health department, and local court.

The Duluth project and model have also produced an influential domestic violence intervention and training program that has taken root in many other states of the union including California. The designers of this program make a number of important assumptions about domestic abuse including the view that most forms of domestic violence are male initiated, with the primary tone of this violence being coercive and instrumental in nature. More specifically, it assumes that male initiated violence in a domestic context is designed to control and even subjugate the female partner to a man's will and needs. The model also assumes that most forms of female initiated violence are primarily defensive and/or retaliatory in nature.<sup>1</sup>

A third assumption is that male client referrals to batterer intervention programs are court ordered, which embeds offenders in the justice system including the courts and probation departments. This last assumption is thought to be critical to the intervention programs success, for it assumes an active and fully supportive judiciary, district attorney's office, and probation department who have found ways of working in a well orchestrated effort to charge, prosecute, convict if guilty, subsequently supervise, and treat domestic violence offenders. The success of this effort is thought to be linked to the existence of clear and certain criminal penalties for noncompliance as well as for re-offense<sup>2</sup>.

Batterer intervention programs based on the Duluth Model will include a number of broad programmatic elements that are thought essential to addressing the basic causes of male domestic violence, which are viewed as being rooted in a belief system that creates rights and expectations for batterers in terms of their roles as partners, fathers, and members of their communities<sup>3</sup>. These program elements include an intake session and group orientation that lays out the obligations and expectations of clients over the course of the 27 week program, acquaints the men with specifics of the class process and course curriculum and materials, and administrative tasks. The various aspects of the curriculum are intended to be educational in nature, with the primary objective of the intervention being to move the core beliefs of male batterers away from a dominant sense of male privilege in relation to their partners and children, with coercion and violence as primary mechanism for achieving this end.

---

<sup>1</sup> Pence, E. Batterer programs: Shifting from community collusion to community confrontation. In P.L. Caesar & L.K. Hamberger (Eds.), *Treating men who batter: Theory, practice and programs*. New York: Springer, 1989.

<sup>2</sup> Pence, E. & Shepard, M. An introduction: Developing a coordinated community response. IN E. Pence & M. Sheppard (Eds.), *Coordinating community responses to domestic violence: Lessons from Duluth and beyond*. Thousand Oaks, CA: Sage, 1999.

<sup>3</sup> Pence, E. & Paymar, M. *Education groups for men who batter: The Duluth model*. New York: Springer, 1993.



### The Duluth curriculum.

Of particular relevance for our understanding of the influence of this model are the educational topics and themes the Duluth Model posits as essential to an effective batterer intervention program. The basic assumptions underlying the Duluth educational curriculum are unique both for what they include and exclude. For example, because the Duluth Model assumes that the root cause of male domestic violence is a result of an internalized societal belief system that promoting the rights of males over others it seeks to reeducate clients rather than treat them psycho-therapeutically. Those approaches to intervention that would ascribe domestic violence to unique psychological or psycho-physiological problems of individuals are excluded from the Duluth model and intervention program.

A number of educational and instructional principles characterize the Duluth educational curriculum. They include but are not restricted to the following.

Curriculum content, teaching strategies, and educational themes.

- Curriculum materials presented in an intervention course should require a literacy level consistent with that of participating clients.
- Educational themes or issue must be directly related to the life experiences of clients.
- Expression of the theme or issue is based on images, pictures, and other materials that avoid needless abstraction, and are rooted in the “real moments of the lives” of clients.
- Themes covered in group should be drawn directly from the Power and Control Wheel that Duluth interventionists have developed. These themes are considered the “tools” of the curriculum.

Themes related to domestic abuse include

- Coercion and threats
- Economic abuse
- Emotional abuse
- Intimidation
- Isolation—isolating one’s partner
- Minimizing, blaming one’s partner, and rationalization of one’s abuse
- Using male privilege to achieve dominance
- Using/manipulating children to get at one’s partner

Themes related to equality and non-abusive attitudes and beliefs include:

- Economic partnership
- Honesty and accountability
- Negotiation and fairness
- Non-threatening behavior
- Responsible parenting
- Shared responsibility

## Appendix F: Duluth and Cognitive-Behavioral Models

- Trust and support

The vehicles for conveying these educational themes rely heavily on the use of video vignettes, group discussion, role playing, and what has been termed a control log. These vehicles allow for communicating complex topics through the use of images and words, and as such are not heavily dependent on the literacy level of group participants<sup>4</sup>.

**The discussion group.** The group discussion and role plays occurring under the careful supervision of the facilitator, are intended to help clients develop the capacity and propensity for critical thinking and accountability. The group becomes one of the primary locations where new ideas are taught and the capacity to think critically is acquired. The environment of the group then must be fully supportive of batterers in their efforts to transform how they think about and behave toward their domestic partners. This then requires that groups hold the abuser fully accountable for his use of violence, while creating an environment that is free of the threat of violence and coercion. It would also have to find a way to be non-judgmental in relation with its members while enforcing a norm of full accountability, as well as requiring participants to be respectful of one another as well as women and children during the course of their group work. Further, the group must require that clients are committed to a lengthy process in which they are deeply honest with themselves and members of the group when discussing their own lives and behavior, while working toward full accountability to the woman they have harmed<sup>5</sup>.

**The group facilitator.** The group facilitator has an important and demanding role in group discussions. This involves guiding participants through the vignettes and challenging them to think critically during the iterative phases of analysis, problem solving, planning, more critical analysis and reflection, and so on. This pedagogical approach allows participants to review and critique abusive behavior that may be highly similar to their own, without having to immediately speak about the specifics of their own abusive attitudes and behavior. This appears to facilitate a deeper form of analysis and reflection than batterers might achieve if they were immediately asked to describe their own behavior, and may help break down the defense mechanisms of denial, minimization, and blame that are often employed by male batterers to thwart challenges to their abusive behavior by others<sup>6</sup>.

**The video vignette.** In order to facilitate their identification with the video simulations, vignettes are developed to reflect the characteristics of group participants, the domestic conditions in which they live, and issues that plausibly portray interchanges between abusive men like them and their partners. The responsibility of group participants, under the careful guidance of the facilitator, is to enter into a conversation with each other that

---

<sup>4</sup> Pence, E. The Duluth domestic abuse intervention project. In E. Aldarondo & F. Mederos (Eds.): Programs for men who batter: Intervention and prevention strategies in a diverse society. Civic Research Institute, 2002.

<sup>5</sup> Pence, E. & Paymar, M. Education groups for men who batter: The Duluth model. New York: Springer, 1993.

<sup>6</sup> Ibid.

## Appendix F: Duluth and Cognitive-Behavioral Models

cycles through analyses of abusive situations. During these conversations group members reflect and comment on what they have seen in each video, formulate alternative non-abusive forms of language and behavior for the male batterers in the film to undertake, engage in further reflective processes on what they have seen and experienced, and further revise their thinking and action plans for participants in the video vignettes, and then return to the reflection and comment<sup>7</sup>.

**The control log.** Within the Duluth framework the power of the structured group discussion is complemented and enhanced by the requirement of a “control log”. Control logs are used to help male clients identify and define their abusive actions and intentions; identify defense mechanisms that help preserve their abusive attitudes and behavior including denial, blaming, and minimization; deconstruct and further analyze “micro-actions” associated with their abuse; identify personal beliefs about the nature of authority, dependency, weakness, self-protection, strength, and love for critical analysis in group; and identify alternatives to specific abuse interactions that the client has engaged in during past relationships<sup>8</sup>.

### The Cognitive-Behavioral Model.

Like the local community approach to domestic violence that would eventually give rise to the Duluth Model, the history of the development of cognitive-behavioral models in the treatment of domestic violence appear to have emerged from the needs of communities to respond to various forms of domestic violence occurring within their boundaries. Activists leading these efforts appear to have first sought out information to advise their intervention efforts, and in doing so discovered few models that would readily direct their efforts to develop intervention programs<sup>9</sup>. Turning to those members of their social networks with expertise in complementary areas of social and clinical practice, they developed intervention programs for batterers that were cognitive-behavioral in their immediate focus, while preserving a response that was more systemic in nature.

Cognitive-behavioral approaches to intervention with domestic batterers appear to embody a number of the same assumptions and imperatives seen in the Duluth model. For example, Dr. Kevin Hamberger’s description of a pioneering application of this approach in Wisconsin states that violence enacted within a domestic relationship is unequivocally unacceptable, and that the objective of domestic abuse is usually the control and domination of one’s domestic partner for self serving purposes. Further, both social and political factors are thought to foster and even facilitate the occurrence and perpetuation of domestic abuse, and that effective responses to such violence necessitate an active collaboration among community agencies and groups. Only then is it thought

---

<sup>7</sup> Ibid.

<sup>8</sup> Ibid.

<sup>9</sup> Hamberger, Kevin. The men’s group program: A community-based, cognitive-behavioral, pro-feminist intervention program. In E. Aldarondo & F. Mederos (Eds.): Programs for men who batter: Intervention and prevention strategies in a diverse society. Civic Research Institute, 2002.

## Appendix F: Duluth and Cognitive-Behavioral Models

that effective advocacy can be developed to change local institutional practices and social norms that place women at risk for violence, and/or fail to hold male batterers responsible for their violent behavior<sup>10</sup>.

While the responsibility for domestic violence is seen to reside solely with the male perpetrator, the violence itself is seen as learned behavior that occurs within a social and cultural context that often sanctions and sometimes tacitly encourages violence against women. Sources of this learning were assumed to include society and culture, the perpetrator's family of origin, as well as various forms of "trial and error learning" that the perpetrator engages in over time<sup>11</sup>. In fact, it is the focus on social learning as the proximal mechanism for domestic violence that may uniquely identify cognitive-behavioral approaches to domestic violence intervention.

Two important conceptual assumptions in this approach are that cognitive processes of labeling and interpretation of life events are related to learned emotional and behavioral responses to similar prior events. The behavior that results is learned and organized through the receipt of reinforcement, with associated cognitive processes linked to these behaviors also receiving indirect reinforcement, with the latter form of reinforcement potentially generalizing an aggressive response from one domain to others quite quickly<sup>12</sup>.

Consequently, a cognitive approach to intervention in domestic violence involves helping batterers understand how they habitually label certain situations as threatening, intolerable, and/or dangerous; the highly negative attributions they make to their domestic partner in these situations; the aggressive coping responses they resort to in order to defend themselves against perceived threats, and the underlying beliefs and more specific cognitive schemata that filter and maintain their existing patterns of thought, which in turn support their dysfunctional behavior. Cognitive behavioral approaches must then address the irrational elements of these cognitive processes, and through a structured and progressive approach to learning, train batterers in new ways of thinking and behaving in relation to their partners.

Within the context of group practice cognitive behavioral approaches may involve a relatively structured set of skill training lessons and exercises. In this approach facilitators are challenged to develop ways of addressing the particular needs and skill deficits of individual batterers within the context of a group. The structure that is required in group sessions involves establishing objectives for each session that involves goal setting, specific active-learning tasks, and criterion-based outcomes to assess and inform group members about the progress they have made. More specifically, early

---

<sup>10</sup> Ibid.

<sup>11</sup> Hamberger, L.K. Cognitive behavioral treatment of men who batter their partners. *Cognitive and Behavioral Practice*, 4, 147-169, 1997.

<sup>12</sup> Hamberger, L.K & Lorch, J.M. Proximal causes of spouse abuse: A theoretical analysis for cognitive-behavioral interventions. In P.L. Caesar & L.K. Hamberger (Eds.), *Treating men who batter: Theory, practice and programs*. New York: Springer (1989).

## Appendix F: Duluth and Cognitive-Behavioral Models

cognitive behavioral approaches are described as involving skills training in the following areas<sup>13</sup>:

- Arousal management which requires the batterer to learn new coping strategies that will allow him to more effectively handle stressful situations in his life, as well as relaxation training intended to augment and support these positive forms of coping.
- Assertive behavior training that will allow for respectful forms of communication with domestic partners, as well skills training in positive forms of conflict resolution.
- Cognitive restructuring including thought switching, with the goal involving the identification of and change in negative labeling and attribution processes providing the basis for the batterers' domestic abuse.

The collaborative approach that emerges involves the group facilitator playing an active leadership role in listening, confronting, coaching, and providing feedback, with more advanced members of the group participating in support of this process with the permission of the facilitator. Clients bring situation-relevant material to session where they actively rehearse alternative thought processes of varying types; and develop multiple options for labeling, interpreting, and self-instructing to deal with threatening domestic situations. The immediate goal is to develop and test specific behavioral strategies in relation to problem behavior or thinking processes that lead the batterer into abuse.

While the cognitive-behavioral approach is typically associated with single client treatment by a highly trained therapist, the model has been adapted to use in groups and appears to share much in common with the Duluth model at this level of operationalization. The Men's Group of Wisconsin employed related the following outline for a typical group session using a cognitive behavioral technique:

1st hour involves 10 minutes of introduction to new members joining the group; with the remaining 50 minutes typically devoted to 1) men sharing coping successes and difficulties, 2) group discussion and analysis including homework, and 3) feedback by the group facilitator.

2nd hour involves

1. didactic lecture and structured discussion that may include:

- beliefs and attitudes associated with male gender role training and its links to DV
- cognitive basis for feelings of jealousy
- development of coping strategies
- outline of personal plans for relating as an equal partner

---

<sup>13</sup> Hamberger, Kevin. The men's group program: A community-based, cognitive-behavioral, pro-feminist intervention program. . In E. Aldarondo & F. Mederos (Eds.): Programs for men who batter: Intervention and prevention strategies in a diverse society. Civic Research Institute, 2002.

## **Appendix F: Duluth and Cognitive-Behavioral Models**

- power and control in relationships,
- role of prejudice and stereotyping in facilitating violence
- understanding the impact of violence on children
- 2. Modeling by therapist or advanced student of new cognitive-behavioral strategy or skill
- 3. Rehearsal of a cognitive skill area by students who have proper readiness
- 4. Consideration of new cognitive-behavioral concept.

It should be clear from a review of this list of activities and training tasks that while this intervention model draws heavily on the cognitive-behavioral tradition, aspects of other intervention models including Duluth are present and even central to this type of intervention in groups.

## Appendix G: Table 3-A1

| Table 3-A1. Program Content Survey: Importance of Educational Topic |                                                                                |                                   |                          |                    |                                       |     |     |     |     |       |
|---------------------------------------------------------------------|--------------------------------------------------------------------------------|-----------------------------------|--------------------------|--------------------|---------------------------------------|-----|-----|-----|-----|-------|
| Item                                                                | Educational Topics Explained or Discussed                                      |                                   |                          | Average Importance | Average Rating of Importance x County |     |     |     |     |       |
|                                                                     | Topic or Issue                                                                 | Coding Cat.                       | N of BIPs Covering Topic |                    | LA                                    | RS  | SC  | SJ  | SoI | Sig.  |
| Q26A                                                                | Gender equality between partners and its implications for everyday behavior.   | Attitudes & Beliefs               | 42                       | 3.9                | 4.3                                   | 3.6 | 3.8 | 3.2 | 3.5 | ns*   |
| Q48A                                                                | Stress and stress management explained.                                        | Stress & Coping                   | 42                       | 3.9                | 4.2                                   | 4.1 | 2.9 | 3.8 | 4.8 | ns    |
| Q3A                                                                 | Alcohol and substance abuse.                                                   | Substance Abuse                   | 43                       | 3.9                | 4.2                                   | 4.6 | 2.8 | 4.0 | 3.5 | ns    |
| Q18A                                                                | Cycle of violence.                                                             | Abuse                             | 42                       | 3.9                | 3.9                                   | 4.1 | 3.0 | 4.7 | 4.8 | ns    |
| Q25A                                                                | Empathy as an essential aspect of close relationships.                         | Empathy                           | 41                       | 3.9                | 3.9                                   | 4.1 | 3.6 | 3.8 | 4.0 | ns    |
| Q52A                                                                | Wheel of Non-Violence in relation to interpersonal relationships.              | Abuse                             | 40                       | 3.9                | 4.2                                   | 4.1 | 3.3 | 2.8 | 4.8 | ns*   |
| Q5A                                                                 | Assertiveness and assertive behavior vs. aggression.                           | Conflict Resolution & Negotiation | 43                       | 3.8                | 4.1                                   | 4.1 | 3.0 | 3.5 | 3.3 | ns    |
| Q34A                                                                | Male privilege and patriarchy as contributor to clients' attitudes & behavior. | Attitudes & Beliefs               | 42                       | 3.8                | 4.0                                   | 3.7 | 3.6 | 2.8 | 4.3 | ns    |
| Q51A                                                                | Violence prevention plan for client.                                           | Planning                          | 37                       | 3.8                | 4.2                                   | 4.0 | 3.1 | 2.3 | 5.0 | ns*   |
| Q37A                                                                | Parenting: Appropriate discipline and punishment of children.                  | Parenting                         | 42                       | 3.7                | 4.1                                   | 3.7 | 2.9 | 3.5 | 3.8 | ns    |
| Q11A                                                                | Client's family as a source of attitudes, beliefs, and abusive behavior.       | Attitudes & Beliefs               | 41                       | 3.6                | 4.2                                   | 4.0 | 3.1 | 2.7 | 2.8 | p<.02 |
| Q28A                                                                | Health vs. unhealthy relationship with partner described in detail.            | Interpersonal health              | 39                       | 3.6                | 4.3                                   | 4.0 | 3.0 | 2.0 | 4.0 | p<.02 |
| Q40A                                                                | Personal boundaries (and the lack thereof) as central to domestic abuse.       | Batterer Characteristics          | 40                       | 3.6                | 4.0                                   | 4.1 | 3.1 | 2.0 | 4.8 | ns*   |
| Q33A                                                                | Jealousy and coping with jealousy.                                             | Anger & Emotion Management        | 41                       | 3.6                | 3.9                                   | 3.6 | 2.6 | 3.7 | 4.5 | ns    |
| Q46A                                                                | Sex role beliefs & expectations as they are related to abuse.                  | Attitudes & Beliefs               | 43                       | 3.6                | 3.8                                   | 4.0 | 3.4 | 3.0 | 3.8 | ns    |
| Q31A                                                                | Interpersonal communication principles & skills explained.                     | Interpersonal communication       | 41                       | 3.6                | 3.8                                   | 4.1 | 3.3 | 3.0 | 3.5 | ns    |
| Q36A                                                                | Negative self talk.                                                            | Cognitive-Behavioral              | 41                       | 3.5                | 4.1                                   | 3.7 | 3.1 | 2.0 | 3.8 | ns*   |
| Q16A                                                                | Cultural and societal norms supporting aggression against women & others.      | Attitudes & Beliefs               | 43                       | 3.5                | 3.8                                   | 3.0 | 3.5 | 3.3 | 2.5 | p<.05 |
| Q10A                                                                | Client's family history of domestic abuse.                                     | Abuse                             | 42                       | 3.4                | 4.2                                   | 4.0 | 3.1 | 2.7 | 2.8 | ns    |
| Q38A                                                                | Parenting: Effective co-parenting.                                             | Parenting                         | 40                       | 3.4                | 4.1                                   | 3.3 | 3.0 | 2.3 | 2.8 | ns*   |
| Q27A                                                                | Handling criticism from spouse or partner.                                     | Stress & Coping                   | 40                       | 3.3                | 3.9                                   | 2.6 | 3.1 | 2.5 | 3.3 | ns*   |
| Q24A                                                                | Emotional sensitization techniques explained.                                  | Stress & Coping                   | 38                       | 3.3                | 3.8                                   | 2.4 | 2.9 | 2.2 | 1.5 | ns*   |
| Q39A                                                                | Parenting: Information, attitudes, and strategies for effective parenting.     | Parenting                         | 39                       | 3.3                | 4.1                                   | 3.4 | 2.6 | 1.7 | 2.8 | ns*   |
| Q9A                                                                 | Characteristics of male batterers.                                             | Batterer Characteristics          | 40                       | 3.3                | 3.9                                   | 3.9 | 2.8 | 2.7 | 3.0 | ns    |
| Q15A                                                                | Coping with separation and/or divorce from partner.                            | Stress & Coping                   | 40                       | 3.2                | 4.0                                   | 3.3 | 2.6 | 1.2 | 3.3 | p<.01 |
| Q19A                                                                | Domestic Abuse: What is it legally.                                            | Abuse                             | 41                       | 3.2                | 3.9                                   | 3.1 | 2.8 | 2.3 | 2.5 | ns    |
| Q47A                                                                | Sexism and sexist oppression.                                                  | Attitudes & Beliefs               | 40                       | 3.2                | 3.5                                   | 3.0 | 3.3 | 2.3 | 3.0 | ns    |
| Q32A                                                                | Interpersonal mis-communication explained.                                     | Interpersonal communication       | 37                       | 3.1                | 3.5                                   | 3.3 | 2.5 | 2.2 | 3.8 | ns    |
| Q23A                                                                | Effects of domestic abuse on other adults & the community.                     | Abuse                             | 38                       | 3.0                | 3.8                                   | 2.4 | 2.9 | 2.2 | 1.5 | ns*   |

## Appendix G: Table 3-A1

| Educational Topics Explained or Discussed |                                                                                       |                          |                          | Average Rating of Importance x County |     |     |     |     |     |       |
|-------------------------------------------|---------------------------------------------------------------------------------------|--------------------------|--------------------------|---------------------------------------|-----|-----|-----|-----|-----|-------|
| Item                                      | Topic or Issue                                                                        | Coding Cat.              | N of BIPs Covering Topic | Average Importance                    | LA  | RS  | SC  | SJ  | SoI | Sig.  |
| Q8A                                       | Victims: Characteristics of abused women (e.g., attitudes, beliefs, & socialization). | Abuse                    | 39                       | 3.0                                   | 3.8 | 3.1 | 2.8 | 2.5 | 3.5 | p<.02 |
| Q12A                                      | Co-Dependency with partner.                                                           | Batterer Characteristics | 36                       | 2.9                                   | 3.9 | 3.0 | 2.0 | 1.0 | 2.5 | ns*   |
| Q44A                                      | Racism as related to clients self concept and attitudes to self and partner.          | Attitudes & Beliefs      | 36                       | 2.8                                   | 3.5 | 2.1 | 2.6 | 2.0 | 1.5 | p<.03 |
| Q42A                                      | Personality disorders and DV.                                                         | Batterer Characteristics | 30                       | 2.3                                   | 2.3 | 2.1 | 2.5 | 1.7 | 2.5 | ns    |
| Q45A                                      | Safety Plan for victim.                                                               | Planning                 | 27                       | 2.3                                   | 2.9 | 2.7 | 1.0 | 1.5 | 2.0 | ns    |



## Appendix H. Table 3-B1

| Table 3B1. Program Content Survey: Frequency of Use of Educational Topics |                                                                                    |                                   |                          |                               |                                    |      |      |      |      |      |
|---------------------------------------------------------------------------|------------------------------------------------------------------------------------|-----------------------------------|--------------------------|-------------------------------|------------------------------------|------|------|------|------|------|
| Item                                                                      | Educational Topics Explained or Discussed                                          |                                   |                          | Average Intensity of Coverage | Relative Intensity of Use x County |      |      |      |      |      |
|                                                                           | Topic or Issue                                                                     | Coding Cat.                       | N of BIPs Covering Topic |                               | LA                                 | RS   | SC   | SJ   | SoI  | Sig. |
| Q53A                                                                      | Wheel of Power & Control in relation to domestic abuse.                            | Power & Control                   | 42                       | 29.2                          | 29.2                               | 28.6 | 33.2 | 25.5 | 28.1 | ns   |
| Q22A                                                                      | Effects of abuse on partner.                                                       | Abuse                             | 43                       | 28.9                          | 25.2                               | 29.9 | 35.9 | 32.7 | 28.1 | ns   |
| Q4A                                                                       | Anger and anger-triggers.                                                          | Anger & Emotion Management        | 43                       | 28.6                          | 27.6                               | 36.1 | 24.1 | 30.7 | 25.6 | ns   |
| Q41A                                                                      | Personal responsibility & honesty on an everyday basis.                            | Accountability                    | 42                       | 28.0                          | 24.5                               | 28.7 | 39.0 | 27.7 | 25.6 | ns   |
| Q21A                                                                      | Effects of abuse on children.                                                      | Abuse                             | 43                       | 27.9                          | 25.6                               | 25.6 | 37.5 | 27.4 | 28.1 | ns   |
| Q18A                                                                      | Cycle of violence.                                                                 | Abuse                             | 42                       | 27.5                          | 25.2                               | 35.8 | 22.5 | 32.6 | 25.8 | ns*  |
| Q14A                                                                      | Conflict resolution techniques.                                                    | Conflict Resolution & Negotiation | 43                       | 27.5                          | 28.6                               | 33.1 | 19.6 | 25.5 | 28.1 | ns   |
| Q13A                                                                      | Cognitive restructuring.                                                           | Cognitive-Behavioral              | 42                       | 27.5                          | 25.1                               | 34.4 | 27.2 | 27.7 | 27.9 | ns   |
| Q30A                                                                      | Identification of high-risk situations.                                            | Stress & Coping                   | 42                       | 27.0                          | 27.6                               | 30.0 | 21.1 | 27.6 | 28.4 | ns   |
| Q26A                                                                      | Gender equality between partners and its implications for everyday behavior.       | Attitudes & Beliefs               | 42                       | 26.3                          | 27.1                               | 19.9 | 34.4 | 23.5 | 22.9 | ns   |
| Q52A                                                                      | Wheel of Non-Violence in relation to interpersonal relationships.                  | Abuse                             | 40                       | 26.2                          | 24.7                               | 22.7 | 31.0 | 30.6 | 28.1 | ns   |
| Q25A                                                                      | Empathy as an essential aspect of close relationships.                             | Empathy                           | 41                       | 26.1                          | 20.9                               | 25.6 | 37.5 | 31.8 | 25.8 | ns   |
| Q6A                                                                       | Beliefs and attitudes leading to domestic abuse.                                   | Attitudes & Beliefs               | 43                       | 25.4                          | 22.1                               | 26.9 | 36.1 | 27.4 | 18.0 | ns   |
| Q28A                                                                      | Health vs. unhealthy relationship with partner described in detail.                | Interpersonal health              | 39                       | 24.4                          | 25.2                               | 25.6 | 29.1 | 17.9 | 18.0 | ns   |
| Q48A                                                                      | Stress and stress management explained.                                            | Stress & Coping                   | 42                       | 24.3                          | 20.3                               | 29.9 | 19.7 | 32.0 | 33.4 | ns   |
| Q40A                                                                      | Personal boundaries (and the lack thereof) as central to domestic abuse.           | Batterer Characteristics          | 40                       | 24.1                          | 22.2                               | 31.4 | 19.7 | 25.5 | 28.1 | ns   |
| Q34A                                                                      | Male privilege and patriarchy as contributor to clients' attitudes & behavior.     | Attitudes & Beliefs               | 42                       | 23.7                          | 20.2                               | 24.1 | 31.4 | 21.6 | 30.9 | ns   |
| Q31A                                                                      | Interpersonal communication principles & skills explained.                         | Interpersonal communication       | 41                       | 23.2                          | 22.6                               | 30.0 | 25.6 | 15.3 | 20.5 | ns   |
| Q1A                                                                       | Accepting and working with victims anger, resentment, distrust as result of abuse. | Accountability                    | 41                       | 22.3                          | 19.3                               | 24.0 | 28.6 | 27.9 | 18.0 | ns   |
| Q32A                                                                      | Interpersonal mis-communication explained.                                         | Interpersonal communication       | 37                       | 21.4                          | 20.2                               | 27.3 | 20.5 | 15.3 | 25.6 | ns   |
| Q36A                                                                      | Negative self talk.                                                                | Cognitive-Behavioral              | 41                       | 21.1                          | 20.7                               | 26.9 | 24.1 | 7.6  | 20.6 | ns*  |
| Q33A                                                                      | Jealousy and coping with jealousy.                                                 | Anger & Emotion Management        | 41                       | 19.4                          | 15.8                               | 21.3 | 18.7 | 27.6 | 25.4 | ns   |
| Q5A                                                                       | Assertiveness and assertive behavior vs. aggression.                               | Conflict Resolution & Negotiation | 43                       | 19.2                          | 21.2                               | 24.1 | 18.1 | 10.2 | 15.4 | ns   |
| Q46A                                                                      | Sex role beliefs & expectations as they are related to abuse.                      | Attitudes & Beliefs               | 43                       | 19.0                          | 21.3                               | 16.8 | 19.8 | 11.9 | 20.5 | ns   |
| Q27A                                                                      | Handling criticism from spouse or partner.                                         | Stress & Coping                   | 40                       | 19.0                          | 19.3                               | 21.4 | 22.6 | 11.2 | 17.9 | ns   |
| Q24A                                                                      | Emotional sensitization techniques explained.                                      | Stress & Coping                   | 38                       | 18.9                          | 16.3                               | 21.4 | 20.4 | 25.6 | 20.4 | ns   |

## Appendix H. Table 3-B1

| Table 3-B1. Program Content Survey: Frequency of Use of Educational Topics |                                                                              |                          |                    |                      |                                    |      |      |      |      |      |
|----------------------------------------------------------------------------|------------------------------------------------------------------------------|--------------------------|--------------------|----------------------|------------------------------------|------|------|------|------|------|
| Item                                                                       | Educational Topics Explained or Discussed                                    |                          |                    | Average Intensity of | Relative Intensity of Use x County |      |      |      |      |      |
|                                                                            | Topic or Issue                                                               | Coding Cat.              | N of BIPs Covering |                      | LA                                 | RS   | SC   | SJ   | SoI  | Sig. |
| Q11A                                                                       | Client's family as a source of attitudes, beliefs, and abusive behavior.     | Attitudes & Beliefs      | 41                 | 18.5                 | 17.3                               | 22.6 | 24.2 | 11.3 | 15.2 | ns   |
| Q3A                                                                        | Alcohol and substance abuse.                                                 | Substance Abuse          | 43                 | 18.3                 | 18.2                               | 27.0 | 18.4 | 13.6 | 10.1 | ns   |
| Q47A                                                                       | Sexism and sexist oppression.                                                | Attitudes & Beliefs      | 40                 | 17.8                 | 19.0                               | 15.4 | 19.7 | 15.2 | 15.4 | ns   |
| Q23A                                                                       | Effects of domestic abuse on other adults & the community.                   | Abuse                    | 38                 | 17.4                 | 16.3                               | 15.3 | 24.2 | 15.3 | 15.3 | ns   |
| Q16A                                                                       | Cultural and societal norms supporting aggression against women & others.    | Attitudes & Beliefs      | 43                 | 16.5                 | 16.3                               | 13.9 | 25.6 | 13.6 | 10.1 | ns   |
| Q15A                                                                       | Coping with separation and/or divorce from partner.                          | Stress & Coping          | 40                 | 16.4                 | 20.1                               | 19.6 | 10.9 | 5.0  | 12.8 | ns*  |
| Q9A                                                                        | Characteristics of male batterers.                                           | Batterer Characteristics | 40                 | 16.3                 | 16.4                               | 19.6 | 15.4 | 17.8 | 10.1 | ns   |
| Q12A                                                                       | Co-Dependency with partner.                                                  | Batterer Characteristics | 36                 | 16.1                 | 18.3                               | 13.9 | 11.1 | 10.3 | 18.8 | ns   |
| Q19A                                                                       | Domestic Abuse: What is it legally.                                          | Abuse                    | 41                 | 16.1                 | 19.0                               | 18.3 | 10.9 | 11.2 | 12.6 | ns   |
| Q37A                                                                       | Parenting: Appropriate discipline and punishment of children.                | Parenting                | 42                 | 15.8                 | 16.9                               | 19.7 | 13.8 | 11.8 | 12.8 | ns   |
| Q45A                                                                       | Safety Plan for victim.                                                      | Planning                 | 27                 | 15.7                 | 21.5                               | 17.5 | 5.0  | 5.0  | 5.0  | ns*  |
| Q38A                                                                       | Parenting: Effective co-parenting.                                           | Parenting                | 40                 | 15.0                 | 15.3                               | 15.2 | 13.8 | 15.3 | 15.3 | ns   |
| Q39A                                                                       | Parenting: Information, attitudes, and strategies for effective parenting.   | Parenting                | 39                 | 14.8                 | 14.3                               | 16.7 | 12.4 | 18.7 | 15.3 | ns   |
| Q44A                                                                       | Racism as related to clients self concept and attitudes to self and partner. | Attitudes & Beliefs      | 36                 | 14.8                 | 19.7                               | 11.3 | 10.9 | 10.1 | 5.0  | ns   |
| Q10A                                                                       | Client's family history of domestic abuse.                                   | Abuse                    | 42                 | 13.9                 | 13.4                               | 18.1 | 13.8 | 13.2 | 10.1 | ns   |
| Q42A                                                                       | Personality disorders and DV.                                                | Batterer Characteristics | 30                 | 12.8                 | 12.4                               | 7.1  | 15.3 | 22.3 | 10.3 | ns   |
| Q8A                                                                        | women (e.g., attitudes, beliefs, & socialization).                           | Abuse                    | 39                 | 11.3                 | 12.9                               | 13.8 | 8.4  | 8.5  | 5.0  | ns   |

## Appendix I. Table 3-C1

| Table 3-C1. Program Content Survey: Importance of Coping Skills Training |                                                                                                                                  |                                   |                          |                    |                                             |     |     |     |     |      |
|--------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|-----------------------------------|--------------------------|--------------------|---------------------------------------------|-----|-----|-----|-----|------|
| Item                                                                     | Coping Skills Training                                                                                                           |                                   |                          |                    | Average Rating of Importance x Jurisdiction |     |     |     |     |      |
|                                                                          | Strategy or Technique                                                                                                            | Coding Cat.                       | N of BIPs Covering Topic | Average Importance | LA                                          | RS  | SC  | SJ  | SoI | Sig. |
| Q1b.                                                                     | Anger management skills & techniques.                                                                                            | Anger & Emotion Management        | 45                       | 4.6                | 4.5                                         | 5.0 | 4.3 | 4.8 | 5.0 | ns*  |
| Q21b.                                                                    | Time-Out technique training & practice.                                                                                          | Anger & Emotion Management        | 44                       | 4.4                | 4.5                                         | 4.7 | 3.4 | 4.5 | 5.0 | ns*  |
| Q5b.                                                                     | Conflict resolution skills and/or techniques.                                                                                    | Conflict Resolution & Negotiation | 45                       | 4.2                | 4.3                                         | 4.6 | 3.6 | 4.0 | 4.8 | ns*  |
| Q4b.                                                                     | Cognitive restructuring techniques to manage negative moods and negative self talk.                                              | Cognitive-Behavioral              | 45                       | 4.2                | 4.3                                         | 4.4 | 3.7 | 3.8 | 4.8 | ns*  |
| Q2b.                                                                     | Assertiveness training (while demonstrating respect for self and partner) as alternative to aggression.                          | Interpersonal Skills              | 44                       | 4.0                | 4.3                                         | 4.0 | 3.1 | 3.7 | 4.5 | ns   |
| Q3b.                                                                     | Client practices analyzing his own behavior to identify the specifics of his abusive style and areas of personal responsibility. | Cognitive-Behavioral (Duluth)     | 41                       | 3.9                | 3.9                                         | 3.4 | 4.4 | 3.7 | 4.3 | ns   |
| Q9b.                                                                     | Emotional expression skills training.                                                                                            | Interpersonal Skills              | 44                       | 3.8                | 3.9                                         | 4.0 | 3.6 | 3.0 | 4.0 | ns   |
| Q7b.                                                                     | Critical thinking skills for clients/abusers.                                                                                    | Cognitive-Behavioral (Duluth )    | 41                       | 3.7                | 4.1                                         | 3.4 | 3.7 | 2.5 | 4.3 | ns*  |
| Q14b.                                                                    | Personal self-control techniques when parenting to avoid abusive behavior.                                                       | Stress & Coping                   | 41                       | 3.7                | 3.8                                         | 4.3 | 3.0 | 3.7 | 3.3 | ns   |
| Q15b.                                                                    | Positive self-talk training.                                                                                                     | Cognitive-Behavioral              | 40                       | 3.7                | 3.9                                         | 3.9 | 3.1 | 2.8 | 4.8 | ns   |
| Q11b.                                                                    | Alternative reactions to perceived problems or threats taught and practiced.                                                     | Cognitive-Behavioral (Duluth)     | 41                       | 3.7                | 3.7                                         | 4.6 | 3.4 | 2.7 | 4.0 | ns   |
| Q19b.                                                                    | Relaxation & stress management training.                                                                                         | Stress & Coping                   | 41                       | 3.6                | 3.8                                         | 3.4 | 3.0 | 3.7 | 4.0 | ns   |
| Q13b.                                                                    | Negotiation and compromise skills training.                                                                                      | Conflict Resolution & Negotiation | 43                       | 3.5                | 3.7                                         | 3.7 | 3.6 | 2.7 | 3.5 | ns   |
| Q18b.                                                                    | Reflective listening training.                                                                                                   | Interpersonal Communication       | 41                       | 3.4                | 3.6                                         | 3.3 | 2.7 | 3.3 | 4.3 | ns   |
| Q6b.                                                                     | Countering technique for irrational or problematic beliefs.                                                                      | Cognitive-Behavioral              | 37                       | 3.4                | 3.8                                         | 3.3 | 3.1 | 1.7 | 4.5 | ns   |
| Q16b.                                                                    | Problem solving skills training for dealing with everyday living including managing finances, time management, etc.              | Problem Solving & Planning        | 39                       | 3.2                | 3.5                                         | 3.6 | 2.7 | 2.5 | 3.3 | ns   |
| Q10b.                                                                    | Emotional sensitization exercises to help client learn to identify his emotions.                                                 | Stress & Coping                   | 36                       | 3.1                | 3.5                                         | 3.1 | 2.1 | 2.5 | 4.0 | ns   |
| Q20b.                                                                    | Thought-switching and reframing training.                                                                                        | Cognitive-Behavioral              | 34                       | 3.0                | 3.2                                         | 3.7 | 1.7 | 1.8 | 4.5 | ns*  |
| Q8b.                                                                     | Decatastrophizing and depathologizing techniques.                                                                                | Cognitive-Behavioral              | 32                       | 2.7                | 3.0                                         | 2.6 | 1.7 | 2.7 | 3.0 | ns   |
| Q12b.                                                                    | Label shifting or re-labeling training.                                                                                          | Cognitive-Behavioral              | 30                       | 2.4                | 2.6                                         | 2.6 | 1.6 | 1.7 | 3.5 | ns   |
| Q17b.                                                                    | Reattribution skills training.                                                                                                   | Cognitive-Behavioral              | 18                       | 1.5                | 2.0                                         | 1.4 | 1.7 | 0.0 | 0.7 | ns*  |

## Appendix J. Table 3-D1

| Table 3-D1. Frequency of Coping Skills Training: Strategies and Techniques |                                                                                                                                  |                                   |                                |                                     |                                         |      |      |      |      |      |
|----------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|-----------------------------------|--------------------------------|-------------------------------------|-----------------------------------------|------|------|------|------|------|
| Item                                                                       | Coping Skills Training                                                                                                           |                                   | N of BIPs<br>Covering<br>Topic | Average<br>Intensity of<br>Coverage | Average Frequency of Use x Jurisdiction |      |      |      |      |      |
|                                                                            | Strategy or Technique                                                                                                            | Coding Cat.                       |                                |                                     | LA                                      | RS   | SC   | SJ   | SoI  | Sig. |
| Q1b.                                                                       | Anger management skills & techniques.                                                                                            | Anger & Emotion Management        | 45                             | 31.9                                | 26.2                                    | 43.5 | 31.6 | 37.7 | 33.4 | ns   |
| Q21b.                                                                      | Time-Out technique training & practice.                                                                                          | Anger & Emotion Management        | 44                             | 30.8                                | 29.6                                    | 38.9 | 27.5 | 30.7 | 28.1 | ns*  |
| Q5b.                                                                       | Conflict resolution skills and/or techniques.                                                                                    | Conflict Resolution & Negotiation | 45                             | 28.4                                | 27.0                                    | 34.5 | 27.2 | 27.3 | 28.1 | ns   |
| Q4b.                                                                       | Cognitive restructuring techniques to manage negative moods and negative self talk.                                              | Cognitive-Behavioral              | 45                             | 26.3                                | 22.7                                    | 32.9 | 25.7 | 29.1 | 30.6 | ns   |
| Q2b.                                                                       | Assertiveness training (while demonstrating respect for self and partner) as alternative to aggression.                          | Interpersonal Skills              | 44                             | 21.2                                | 21.8                                    | 26.9 | 18.8 | 13.7 | 23.1 | ns   |
| Q3b.                                                                       | Client practices analyzing his own behavior to identify the specifics of his abusive style and areas of personal responsibility. | Cognitive-Behavioral (Duluth)     | 41                             | 29.7                                | 24.6                                    | 42.1 | 31.6 | 34.1 | 30.9 | ns   |
| Q9b.                                                                       | Emotional expression skills training.                                                                                            | Interpersonal Skills              | 44                             | 25.4                                | 24.2                                    | 30.1 | 25.6 | 23.6 | 25.6 | ns   |
| Q7b.                                                                       | Critical thinking skills for clients/abusers.                                                                                    | Cognitive-Behavioral (Duluth )    | 41                             | 24.9                                | 23.2                                    | 23.8 | 34.3 | 20.4 | 25.6 | ns   |
| Q14b.                                                                      | Personal self-control techniques when parenting to avoid abusive behavior.                                                       | Stress & Coping                   | 41                             | 20.9                                | 20.8                                    | 24.2 | 17.0 | 20.6 | 22.3 | ns   |
| Q15b.                                                                      | Positive self-talk training.                                                                                                     | Cognitive-Behavioral              | 40                             | 23.1                                | 23.5                                    | 22.6 | 20.5 | 25.5 | 23.1 | ns   |
| Q11b.                                                                      | Alternative reactions to perceived problems or threats taught and practiced.                                                     | Cognitive-Behavioral (Duluth)     | 41                             | 26.4                                | 23.9                                    | 35.9 | 34.2 | 17.5 | 20.6 | ns   |
| Q19b.                                                                      | Relaxation & stress management training.                                                                                         | Stress & Coping                   | 41                             | 20.6                                | 20.2                                    | 27.3 | 13.8 | 25.8 | 18.0 | ns   |
| Q13b.                                                                      | Negotiation and compromise skills training.                                                                                      | Conflict Resolution & Negotiation | 43                             | 20.3                                | 22.5                                    | 21.1 | 16.7 | 17.4 | 18.0 | ns   |
| Q18b.                                                                      | Reflective listening training.                                                                                                   | Interpersonal Communication       | 41                             | 18.6                                | 20.5                                    | 16.8 | 15.3 | 18.7 | 18.0 | ns   |
| Q6b.                                                                       | Countering technique for irrational or problematic beliefs.                                                                      | Cognitive-Behavioral              | 37                             | 26.2                                | 26.2                                    | 25.6 | 36.1 | 11.8 | 25.6 | ns   |
| Q16b.                                                                      | Problem solving skills training for dealing with everyday living including managing finances, time management, etc.              | Problem Solving & Planning        | 39                             | 19.8                                | 16.9                                    | 28.9 | 15.3 | 30.8 | 15.3 | ns   |
| Q10b.                                                                      | Emotional sensitization exercises to help client learn to identify his emotions.                                                 | Stress & Coping                   | 36                             | 23.7                                | 22.4                                    | 31.8 | 23.3 | 23.1 | 20.6 | ns   |
| Q20b.                                                                      | Thought-switching and reframing training.                                                                                        | Cognitive-Behavioral              | 34                             | 23.3                                | 22.1                                    | 24.1 | 29.2 | 22.3 | 23.1 | ns   |
| Q8b.                                                                       | Decatastrophizing and depathologizing techniques.                                                                                | Cognitive-Behavioral              | 32                             | 22.4                                | 23.0                                    | 28.0 | 17.8 | 19.6 | 22.3 | ns   |
| Q12b.                                                                      | Label shifting or re-labeling training.                                                                                          | Cognitive-Behavioral              | 30                             | 16.4                                | 17.4                                    | 21.6 | 15.2 | 8.5  | 12.8 | ns   |
| Q17b.                                                                      | Reattribution skills training.                                                                                                   | Cognitive-Behavioral              | 18                             | 15.8                                | 14.4                                    | 22.0 | 18.7 | .    | 5.0  | ns   |

## Appendix K. Table 3-E1

| Table 3-E1. Importance and Frequency of Teaching Strategies and Techniques |                                                                                                                                                                      |                          |                    |      |                              |     |     |     |     |       |
|----------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|--------------------|------|------------------------------|-----|-----|-----|-----|-------|
| Item                                                                       | Teaching Techniques and Strategies                                                                                                                                   |                          |                    |      | Average Rating of Importance |     |     |     |     |       |
|                                                                            | Topic or Issue                                                                                                                                                       | N of BIPs Covering Topic | Average Importance |      | LA                           | RS  | SC  | SJ  | Sol | Sig.  |
| 7c.                                                                        | Group discussion: Structured and led by facilitator.                                                                                                                 | 44                       | 4.6                | 41.9 | 4.4                          | 5.0 | 4.6 | 4.7 | 5.0 | ns*   |
| 1c.                                                                        | Client instructed in the analysis of his own abusive behavior to become aware of personal anger triggers and other aspects of his abusive style & cycle of violence. | 43                       | 4.3                | 30.5 | 4.3                          | 4.4 | 3.6 | 4.8 | 5.0 | ns*   |
| 26c.                                                                       | Therapeutic/educational confrontation of clients by group facilitator.                                                                                               | 41                       | 4.0                | 33   | 4.0                          | 3.7 | 3.9 | 3.8 | 4.8 | ns    |
| 9c.                                                                        | Group members allowed to take the lead in challenging attitudes and beliefs that encourage domestic violence.                                                        | 40                       | 3.7                | 29.8 | 3.6                          | 2.9 | 4.7 | 3.3 | 4.0 | ns*   |
| 3c.                                                                        | Facilitator leads clients through a description of some of his most severe incidents of partner abuse.                                                               | 40                       | 3.6                | 23.6 | 3.5                          | 3.1 | 4.3 | 3.7 | 3.8 | ns    |
| 18c.                                                                       | Lecture or formal presentation by facilitator.                                                                                                                       | 37                       | 3.5                | 31.6 | 3.4                          | 3.7 | 3.7 | 3.5 | 3.3 | ns    |
| 10c.                                                                       | Homework: Client develops prevention or safety plan to prevent future abuse.                                                                                         | 38                       | 3.4                | 21.3 | 3.2                          | 3.0 | 3.7 | 3.7 | 4.3 | ns    |
| 24c.                                                                       | Role-playing led by group facilitator.                                                                                                                               | 39                       | 3.3                | 17.5 | 3.8                          | 2.7 | 3.0 | 2.2 | 4.3 | ns    |
| 21c.                                                                       | Rehearsal of cognitive and behavioral skills in group                                                                                                                | 37                       | 3.3                | 24   | 3.5                          | 3.1 | 3.4 | 2.3 | 3.8 | ns    |
| 22c.                                                                       | Rehearsal of coping strategies (e.g. Time-out, etc.).                                                                                                                | 35                       | 3.3                | 26.2 | 3.6                          | 3.3 | 2.7 | 2.0 | 4.5 | ns    |
| 25c.                                                                       | Therapeutic/educational confrontation of clients by "advanced students/clients" in group sessions.                                                                   | 33                       | 3.1                | 25.3 | 3.5                          | 3.0 | 2.9 | 2.2 | 2.8 | ns    |
| 6c.                                                                        | Films & Videos: Developed specifically for domestic violence courses.                                                                                                | 35                       | 2.9                | 11.5 | 2.9                          | 3.1 | 2.9 | 2.0 | 4.3 | ns    |
| 11c.                                                                       | Homework: Client keeps track of the specifics of his abusive behavior and is required to identify areas of personal responsibility & accountability.                 | 31                       | 2.9                |      | 3.1                          | 2.9 | 3.7 | 1.5 | 2.5 | ns    |
| 17c.                                                                       | Homework: Writing assignments based on themes or topics presented in group session.                                                                                  | 33                       | 2.8                |      | 3.0                          | 2.0 | 3.1 | 2.7 | 3.3 | ns*   |
| 19c.                                                                       | Mirroring technique toward validating client's feelings.                                                                                                             | 33                       | 2.8                |      | 3.5                          | 3.3 | 1.9 | 1.2 | 2.3 | ns    |
| 4c.                                                                        | Female facilitators lead groups to address gender-based issues of client trust, identification, and/or attachment.                                                   | 31                       | 2.8                |      | 3.1                          | 3.0 | 3.4 | 0.7 | 2.5 | ns*   |
| 5c.                                                                        | Films & Videos: Not specifically made for DV courses, but relevant to domestic abuse.                                                                                | 38                       | 2.7                |      | 3.5                          | 2.0 | 2.9 | 0.8 | 2.5 | p<.01 |
| 20c.                                                                       | Quizzes and test for checking client's progress & mastery of course materials.                                                                                       | 31                       | 2.6                |      | 2.9                          | 2.0 | 3.7 | 0.3 | 3.0 | ns*   |
| 14c.                                                                       | Homework: Letter of accountability.                                                                                                                                  | 24                       | 2.2                |      | 2.1                          | 1.6 | 2.0 | 2.5 | 3.5 | ns*   |
| 23c.                                                                       | Role-playing led by an advanced group member.                                                                                                                        | 24                       | 2.1                |      | 3.1                          | 1.3 | 2.1 | 0.0 | 1.3 | ns*   |
| 27c.                                                                       | Unstructured group discussions focused on the client's semi-conscious sense of helplessness relative to the partner, fear of abandonment, and/or sense of shame.     | 25                       | 2.0                |      | 2.6                          | 2.0 | 1.9 | 0.7 | 1.8 | ns    |
| 2c.                                                                        | Co-leadership of group by two or more facilitators.                                                                                                                  | 25                       | 2.0                |      | 2.1                          | 1.9 | 4.4 | 0.5 | 0.3 | ns*   |
| 15c.                                                                       | Homework: Reading assignments.                                                                                                                                       | 24                       | 2.0                |      | 1.9                          | 2.1 | 1.4 | 2.2 | 3.3 | ns    |
| 16c.                                                                       | Homework: Relapse prevention plan for client.                                                                                                                        | 21                       | 2.0                |      | 2.2                          | 1.6 | 1.4 | 1.5 | 3.3 | ns    |
| 13c.                                                                       | Homework: Client's controlling behavior log.                                                                                                                         | 22                       | 1.8                |      | 2.1                          | 0.9 | 3.4 | 0.7 | 1.3 | ns    |
| 12c.                                                                       | Homework: Client's anger journal assigned on a regular basis.                                                                                                        | 19                       | 1.8                |      | 2.3                          | 1.3 | 1.1 | 0.8 | 2.5 | ns    |
| 8c.                                                                        | Group discussion: Not structured by facilitator.                                                                                                                     | 17                       | 1.4                |      | 1.8                          | 0.1 | 2.1 | 0.5 | 1.5 | ns    |

# Appendix L. Chapter 5 Regression Tables with Restricted Samples in All Models

**Table L5-C. Regression Results for Program Completion**

| Dependent Variable:<br>Program Completion                                                  | Logistic Regression Models |                   |                   |                   | Multilevel Logistic Regression Models |                   |                   |
|--------------------------------------------------------------------------------------------|----------------------------|-------------------|-------------------|-------------------|---------------------------------------|-------------------|-------------------|
|                                                                                            | (1) <sup>†</sup>           | (2)               | (3)               | (4)               | (5)                                   | (6)               | (7)               |
| <i>Jurisdiction (Los Angeles as the base comparison group)</i>                             |                            |                   |                   |                   |                                       |                   |                   |
| Riverside                                                                                  | 1.213<br>(0.81)            | 1.213<br>(0.83)   | 1.394<br>(1.46)   | 1.498<br>(1.64)   | 1.239<br>(0.61)                       | 1.396<br>(1.04)   | 1.456<br>(1.08)   |
| Santa Clara                                                                                | 0.773<br>(1.48)            | 0.773<br>(1.09)   | 0.888<br>(0.46)   | 0.886<br>(0.47)   | 0.864<br>(0.52)                       | 0.971<br>(0.12)   | 0.958<br>(0.15)   |
| Solano                                                                                     | 2.442<br>(2.07)*           | 2.442<br>(3.11)** | 2.776<br>(2.93)** | 3.709<br>(2.61)** | 2.620<br>(1.83)                       | 2.924<br>(2.15)*  | 3.895<br>(2.54)*  |
| San Joaquin                                                                                | 0.806<br>(1.04)            | 0.806<br>(0.97)   | 1.053<br>(0.23)   | 1.233<br>(0.87)   | 0.859<br>(0.45)                       | 1.073<br>(0.23)   | 1.265<br>(0.71)   |
| <i>No children as base comparison group</i>                                                |                            |                   |                   |                   |                                       |                   |                   |
| Lives with children                                                                        |                            |                   | 1.061<br>(0.30)   | 1.088<br>(0.40)   |                                       | 1.086<br>(0.41)   | 1.115<br>(0.50)   |
| Visits children regularly                                                                  |                            |                   | 0.736<br>(1.52)   | 0.804<br>(0.86)   |                                       | 0.756<br>(1.36)   | 0.829<br>(0.85)   |
| Does not visit children regularly                                                          |                            |                   | 0.892<br>(0.37)   | 0.810<br>(0.59)   |                                       | 0.881<br>(0.51)   | 0.806<br>(0.79)   |
| Education: some college or more                                                            |                            |                   | 1.827<br>(2.81)** | 1.543<br>(1.65)   |                                       | 1.866<br>(3.37)** | 1.574<br>(2.20)*  |
| Victim is wife - former and current                                                        |                            |                   | 1.644<br>(3.05)** | 1.398<br>(1.76)   |                                       | 1.606<br>(2.90)** | 1.353<br>(1.69)   |
| Non-English Speaker                                                                        |                            |                   | 2.015<br>(4.04)** | 1.061<br>(0.32)   |                                       | 2.093<br>(3.65)** | 1.089<br>(0.35)   |
| <i>Employment Status (employed as base comparison group)</i>                               |                            |                   |                   |                   |                                       |                   |                   |
| Employed part-time                                                                         |                            |                   |                   | 1.007<br>(0.03)   |                                       |                   | 0.975<br>(0.11)   |
| Not employed                                                                               |                            |                   |                   | 0.785<br>(1.43)   |                                       |                   | 0.780<br>(1.33)   |
| <i>Race/ethnicity (African American as base comparison group)</i>                          |                            |                   |                   |                   |                                       |                   |                   |
| Hispanic                                                                                   |                            |                   |                   | 2.613<br>(2.93)** |                                       |                   | 2.717<br>(3.96)** |
| Other                                                                                      |                            |                   |                   | 1.926<br>(1.46)   |                                       |                   | 2.014<br>(1.93)   |
| White                                                                                      |                            |                   |                   | 1.629<br>(1.27)   |                                       |                   | 1.664<br>(1.88)   |
| <i>Total prior arrests of any offense (single, first offense as base comparison group)</i> |                            |                   |                   |                   |                                       |                   |                   |
| 2-3                                                                                        |                            |                   |                   | 0.797<br>(0.71)   |                                       |                   | 0.768<br>(0.87)   |
| 4-5                                                                                        |                            |                   |                   | 0.621<br>(1.44)   |                                       |                   | 0.601<br>(1.41)   |
| >=6                                                                                        |                            |                   |                   | 0.551<br>(1.70)   |                                       |                   | 0.540<br>(1.60)   |
| <i>Total prior arrests of DV offenses (single, first offense as base comparison group)</i> |                            |                   |                   |                   |                                       |                   |                   |
| 2                                                                                          |                            |                   |                   | 1.013<br>(0.07)   |                                       |                   | 1.002<br>(0.01)   |
| >=3                                                                                        |                            |                   |                   | 0.760<br>(0.95)   |                                       |                   | 0.751<br>(1.28)   |
| Had prior felony arrests                                                                   |                            |                   |                   | 0.782<br>(0.83)   |                                       |                   | 0.793<br>(0.78)   |
| Had prior arrests for drug offenses                                                        |                            |                   |                   | 0.700<br>(2.58)** |                                       |                   | 0.704<br>(1.79)   |
| Age at Intake                                                                              |                            |                   |                   | 1.036<br>(3.32)** |                                       |                   | 1.036<br>(2.82)** |
| Age at First Arrest                                                                        |                            |                   |                   | 0.999<br>(0.10)   |                                       |                   | 0.998<br>(0.10)   |
| <i>CAGE Score (0 as base comparison group)</i>                                             |                            |                   |                   |                   |                                       |                   |                   |
| CAGE = 1                                                                                   |                            |                   |                   | 0.953<br>(0.16)   |                                       |                   | 0.932<br>(0.30)   |
| CAGE = 2                                                                                   |                            |                   |                   | 0.668<br>(1.88)   |                                       |                   | 0.665<br>(1.78)   |
| CAGE = 3                                                                                   |                            |                   |                   | 0.444<br>(3.54)** |                                       |                   | 0.427<br>(3.27)** |
| CAGE = 4                                                                                   |                            |                   |                   | 0.705<br>(1.17)   |                                       |                   | 0.700<br>(1.29)   |
| Program level variance                                                                     |                            |                   |                   |                   | 0.196<br>(1.28)                       | 0.108<br>(0.96)   | 0.131<br>(1.01)   |
| Observations                                                                               | 802                        | 802               | 802               | 802               | 802                                   | 802               | 802               |

Absolute value of z statistics in parentheses; \* significant at 5%; \*\* significant at 1%.

<sup>†</sup> In models 2 to 4, Huber-White sandwich estimate of variance is used for robust estimates of standard errors that adjust for intraclass correlations at the level of batterer intervention programs.

# Appendix L. Chapter 5 Regression Tables with Restricted Samples in All Models

**Table L5-D. Regression Results for Program Termination**

| Dependent Variable:                                                                        | Logistic Regression Models |                   |                   |                   | Multilevel Logistic Regression Models |                   |                   |
|--------------------------------------------------------------------------------------------|----------------------------|-------------------|-------------------|-------------------|---------------------------------------|-------------------|-------------------|
| Program Termination                                                                        | (1) <sup>†</sup>           | (2)               | (3)               | (4)               | (5)                                   | (6)               | (7)               |
| <i>Jurisdiction (Los Angeles as the base comparison group)</i>                             |                            |                   |                   |                   |                                       |                   |                   |
| Riverside                                                                                  | 0.932<br>(0.32)            | 0.932<br>(0.31)   | 0.775<br>(1.07)   | 0.763<br>(1.09)   | 0.902<br>(0.31)                       | 0.771<br>(0.83)   | 0.778<br>(0.76)   |
| Santa Clara                                                                                | 1.591<br>(2.76)**          | 1.591<br>(2.33)*  | 1.357<br>(1.46)   | 1.344<br>(1.44)   | 1.408<br>(1.30)                       | 1.215<br>(0.78)   | 1.222<br>(0.75)   |
| Solano                                                                                     | 0.523<br>(1.61)            | 0.523<br>(1.78)   | 0.444<br>(1.81)   | 0.357<br>(1.75)   | 0.491<br>(1.46)                       | 0.422<br>(1.81)   | 0.339<br>(2.14)*  |
| San Joaquin                                                                                | 2.504<br>(4.58)**          | 2.504<br>(2.98)** | 1.912<br>(2.07)*  | 1.753<br>(1.88)   | 2.752<br>(3.13)**                     | 2.143<br>(2.50)*  | 1.964<br>(2.08)*  |
| <i>No children as base comparison group</i>                                                |                            |                   |                   |                   |                                       |                   |                   |
| Lives with children                                                                        |                            |                   | 0.815<br>(1.01)   | 0.800<br>(1.08)   |                                       | 0.799<br>(1.13)   | 0.781<br>(1.18)   |
| Visits children regularly                                                                  |                            |                   | 1.148<br>(0.91)   | 1.043<br>(0.21)   |                                       | 1.112<br>(0.53)   | 1.004<br>(0.02)   |
| Does not visit children regularly                                                          |                            |                   | 0.975<br>(0.09)   | 1.068<br>(0.22)   |                                       | 0.987<br>(0.05)   | 1.079<br>(0.29)   |
| Education: some college or more                                                            |                            |                   | 0.531<br>(3.26)** | 0.604<br>(2.27)*  |                                       | 0.511<br>(3.85)** | 0.577<br>(2.88)** |
| Victim is wife - former and current                                                        |                            |                   | 0.618<br>(3.26)** | 0.756<br>(1.63)   |                                       | 0.627<br>(3.00)** | 0.772<br>(1.53)   |
| Non-English Speaker                                                                        |                            |                   | 0.467<br>(4.88)** | 0.812<br>(1.24)   |                                       | 0.459<br>(3.98)** | 0.797<br>(0.99)   |
| <i>Employment Status (employed as base comparison group)</i>                               |                            |                   |                   |                   |                                       |                   |                   |
| Employed part-time                                                                         |                            |                   |                   | 1.094<br>(0.46)   |                                       |                   | 1.109<br>(0.47)   |
| Not employed                                                                               |                            |                   |                   | 1.468<br>(2.23)*  |                                       |                   | 1.451<br>(2.10)*  |
| <i>Race/ethnicity (African American as base comparison group)</i>                          |                            |                   |                   |                   |                                       |                   |                   |
| Hispanic                                                                                   |                            |                   |                   | 0.464<br>(2.41)*  |                                       |                   | 0.441<br>(3.37)** |
| Other                                                                                      |                            |                   |                   | 0.544<br>(1.43)   |                                       |                   | 0.520<br>(1.90)   |
| White                                                                                      |                            |                   |                   | 0.658<br>(1.22)   |                                       |                   | 0.625<br>(1.79)   |
| <i>Total prior arrests of any offense (single, first offense as base comparison group)</i> |                            |                   |                   |                   |                                       |                   |                   |
| 2-3                                                                                        |                            |                   |                   | 1.510<br>(1.57)   |                                       |                   | 1.538<br>(1.55)   |
| 4-5                                                                                        |                            |                   |                   | 2.052<br>(2.00)*  |                                       |                   | 2.023<br>(2.10)*  |
| >=6                                                                                        |                            |                   |                   | 2.089<br>(2.13)*  |                                       |                   | 2.063<br>(1.98)*  |
| <i>Total prior arrests of DV offenses (single, first offense as base comparison group)</i> |                            |                   |                   |                   |                                       |                   |                   |
| 2                                                                                          |                            |                   |                   | 0.960<br>(0.19)   |                                       |                   | 0.963<br>(0.18)   |
| >=3                                                                                        |                            |                   |                   | 1.267<br>(0.95)   |                                       |                   | 1.298<br>(1.20)   |
| Had prior felony arrests                                                                   |                            |                   |                   | 1.068<br>(0.26)   |                                       |                   | 1.091<br>(0.31)   |
| Had prior arrests for drug offenses                                                        |                            |                   |                   | 1.238<br>(1.73)   |                                       |                   | 1.243<br>(1.15)   |
| Age at Intake                                                                              |                            |                   |                   | 0.972<br>(2.52)*  |                                       |                   | 0.971<br>(2.42)*  |
| Age at First Arrest                                                                        |                            |                   |                   | 1.001<br>(0.04)   |                                       |                   | 1.001<br>(0.09)   |
| <i>CAGE Score (0 as base comparison group)</i>                                             |                            |                   |                   |                   |                                       |                   |                   |
| CAGE = 1                                                                                   |                            |                   |                   | 0.990<br>(0.04)   |                                       |                   | 1.023<br>(0.10)   |
| CAGE = 2                                                                                   |                            |                   |                   | 1.532<br>(1.71)   |                                       |                   | 1.534<br>(1.96)   |
| CAGE = 3                                                                                   |                            |                   |                   | 1.820<br>(2.69)** |                                       |                   | 1.891<br>(2.56)*  |
| CAGE = 4                                                                                   |                            |                   |                   | 1.320<br>(0.96)   |                                       |                   | 1.354<br>(1.14)   |
| Program level variance                                                                     |                            |                   |                   |                   | 0.162<br>(1.41)                       | 0.112<br>(1.19)   | 0.128<br>(1.16)   |
| Observations                                                                               | 892                        | 892               | 892               | 892               | 892                                   | 892               | 892               |

Absolute value of z statistics in parentheses; \* significant at 5%; \*\* significant at 1%.

<sup>†</sup> In models 2 to 4, Huber-White sandwich estimate of variance is used for robust estimates of standard errors that adjust for intraclass correlations at the level of batterer intervention programs.

# Appendix L. Chapter 5 Regression Tables with Restricted Samples in All Models

**Table L5-H. Regression Results for 12-Month Re-arrests of All Offense Types**

| Dependent Variable:<br>12-Month Rearrests of All Offense Types                             | Logistic Regression Models |                   |                   |                   | Multilevel Logistic Regression Models |                   |                   |
|--------------------------------------------------------------------------------------------|----------------------------|-------------------|-------------------|-------------------|---------------------------------------|-------------------|-------------------|
|                                                                                            | (1)                        | (2)               | (3)               | (4)               | (5)                                   | (6)               | (7)               |
| <i>Jurisdiction (Los Angeles as the base comparison group)</i>                             |                            |                   |                   |                   |                                       |                   |                   |
| Riverside                                                                                  | 1.188<br>(0.80)            | 1.188<br>(0.87)   | 1.042<br>(0.19)   | 1.112<br>(0.39)   | 1.191<br>(0.79)                       | 1.047<br>(0.20)   | 1.131<br>(0.47)   |
| Santa Clara                                                                                | 1.836<br>(3.68)**          | 1.836<br>(3.12)** | 1.685<br>(2.66)** | 1.683<br>(2.73)** | 1.810<br>(3.31)**                     | 1.657<br>(2.70)** | 1.650<br>(2.40)*  |
| Solano                                                                                     | 0.408<br>(1.91)            | 0.408<br>(6.05)** | 0.375<br>(5.45)** | 0.335<br>(2.93)** | 0.406<br>(1.91)                       | 0.373<br>(2.04)*  | 0.332<br>(2.08)*  |
| San Joaquin                                                                                | 1.221<br>(1.05)            | 1.221<br>(1.16)   | 0.979<br>(0.12)   | 0.865<br>(0.84)   | 1.218<br>(0.99)                       | 0.977<br>(0.11)   | 0.858<br>(0.65)   |
| <i>No children as base comparison group</i>                                                |                            |                   |                   |                   |                                       |                   |                   |
| Lives with children                                                                        |                            |                   | 0.899<br>(0.49)   | 0.929<br>(0.35)   |                                       | 0.898<br>(0.57)   | 0.928<br>(0.37)   |
| Visits children regularly                                                                  |                            |                   | 1.281<br>(1.36)   | 1.240<br>(1.09)   |                                       | 1.281<br>(1.33)   | 1.236<br>(1.05)   |
| Does not visit children regularly                                                          |                            |                   | 0.856<br>(0.56)   | 1.015<br>(0.06)   |                                       | 0.855<br>(0.68)   | 1.013<br>(0.05)   |
| Education: some college or more                                                            |                            |                   | 0.525<br>(4.07)** | 0.632<br>(2.16)*  |                                       | 0.522<br>(3.78)** | 0.627<br>(2.45)*  |
| Victim is wife - former and current                                                        |                            |                   | 0.604<br>(4.01)** | 0.731<br>(2.06)*  |                                       | 0.603<br>(3.40)** | 0.728<br>(1.92)   |
| Non-English Speaker                                                                        |                            |                   | 0.659<br>(2.55)*  | 1.463<br>(1.75)   |                                       | 0.662<br>(2.23)*  | 1.477<br>(1.71)   |
| <i>Employment Status (employed as base comparison group)</i>                               |                            |                   |                   |                   |                                       |                   |                   |
| Employed part-time                                                                         |                            |                   |                   | 1.335<br>(1.15)   |                                       |                   | 1.337<br>(1.34)   |
| Not employed                                                                               |                            |                   |                   | 1.481<br>(2.43)*  |                                       |                   | 1.481<br>(2.29)*  |
| <i>Race/ethnicity (African American as base comparison group)</i>                          |                            |                   |                   |                   |                                       |                   |                   |
| Hispanic                                                                                   |                            |                   |                   | 0.621<br>(2.20)*  |                                       |                   | 0.616<br>(2.16)*  |
| Other                                                                                      |                            |                   |                   | 0.984<br>(0.04)   |                                       |                   | 0.978<br>(0.06)   |
| White                                                                                      |                            |                   |                   | 0.885<br>(0.47)   |                                       |                   | 0.872<br>(0.56)   |
| <i>Total prior arrests of any offense (single, first offense as base comparison group)</i> |                            |                   |                   |                   |                                       |                   |                   |
| 2-3                                                                                        |                            |                   |                   | 3.238<br>(4.36)** |                                       |                   | 3.259<br>(3.84)** |
| 4-5                                                                                        |                            |                   |                   | 2.921<br>(2.94)** |                                       |                   | 2.945<br>(3.04)** |
| >=6                                                                                        |                            |                   |                   | 8.309<br>(5.62)** |                                       |                   | 8.338<br>(5.60)** |
| <i>Total prior arrests of DV offenses (single, first offense as base comparison group)</i> |                            |                   |                   |                   |                                       |                   |                   |
| 2                                                                                          |                            |                   |                   | 0.934<br>(0.36)   |                                       |                   | 0.938<br>(0.33)   |
| >=3                                                                                        |                            |                   |                   | 0.651<br>(1.94)   |                                       |                   | 0.653<br>(2.09)*  |
| Had prior felony arrests                                                                   |                            |                   |                   | 1.428<br>(1.30)   |                                       |                   | 1.419<br>(1.16)   |
| Had prior arrests for drug offenses                                                        |                            |                   |                   | 0.886<br>(0.70)   |                                       |                   | 0.883<br>(0.69)   |
| Age at Intake                                                                              |                            |                   |                   | 0.972<br>(2.48)*  |                                       |                   | 0.972<br>(2.46)*  |
| Age at First Arrest                                                                        |                            |                   |                   | 0.990<br>(0.70)   |                                       |                   | 0.990<br>(0.69)   |
| <i>CAGE Score (0 as base comparison group)</i>                                             |                            |                   |                   |                   |                                       |                   |                   |
| CAGE = 1                                                                                   |                            |                   |                   | 1.005<br>(0.02)   |                                       |                   | 1.014<br>(0.06)   |
| CAGE = 2                                                                                   |                            |                   |                   | 1.572<br>(2.33)*  |                                       |                   | 1.576<br>(2.14)*  |
| CAGE = 3                                                                                   |                            |                   |                   | 1.301<br>(1.03)   |                                       |                   | 1.312<br>(1.15)   |
| CAGE = 4                                                                                   |                            |                   |                   | 1.239<br>(0.88)   |                                       |                   | 1.244<br>(0.86)   |
| Program level variance                                                                     |                            |                   |                   |                   | 0.000<br>(0.00)                       | 0.000<br>(0.00)   | 0.000<br>(0.00)   |
| Observations                                                                               | 941                        | 941               | 941               | 941               | 941                                   | 941               | 941               |

Absolute value of z statistics in parentheses; \* significant at 5%; \*\* significant at 1%.

<sup>†</sup> In models 2 to 4, Huber-White sandwich estimate of variance is used for robust estimates of standard errors that adjust for intraclass correlations at the level of batterer intervention programs.



# Appendix L. Chapter 5 Regression Tables with Restricted Samples in All Models

**Table L5-I. Regression Results for 12-Month Re-arrests of Domestic Violence Offense**

| Dependent Variable:<br>12-Month Rearrests of DV Offense                                    | Logistic Regression Models |                   |                   |                   | Multilevel Logistic Regression Models |                   |                   |
|--------------------------------------------------------------------------------------------|----------------------------|-------------------|-------------------|-------------------|---------------------------------------|-------------------|-------------------|
|                                                                                            | (1)                        | (2)               | (3)               | (4)               | (5)                                   | (6)               | (7)               |
| <i>Jurisdiction (Los Angeles as the base comparison group)</i>                             |                            |                   |                   |                   |                                       |                   |                   |
| Riverside                                                                                  | 1.019<br>(0.07)            | 1.019<br>(0.11)   | 0.811<br>(1.29)   | 0.854<br>(0.85)   | 1.019<br>(0.07)                       | 0.811<br>(0.72)   | 0.854<br>(0.52)   |
| Santa Clara                                                                                | 1.690<br>(2.60)**          | 1.690<br>(3.15)** | 1.538<br>(2.58)** | 1.558<br>(2.48)*  | 1.690<br>(2.60)**                     | 1.538<br>(2.06)*  | 1.558<br>(1.96)*  |
| Solano                                                                                     | 0.741<br>(0.54)            | 0.741<br>(1.50)   | 0.620<br>(1.83)   | 0.484<br>(1.69)   | 0.741<br>(0.54)                       | 0.620<br>(0.85)   | 0.484<br>(1.24)   |
| San Joaquin                                                                                | 1.413<br>(1.47)            | 1.413<br>(1.92)   | 1.067<br>(0.37)   | 0.948<br>(0.25)   | 1.413<br>(1.47)                       | 1.067<br>(0.27)   | 0.948<br>(0.21)   |
| <i>No children as base comparison group</i>                                                |                            |                   |                   |                   |                                       |                   |                   |
| Lives with children                                                                        |                            |                   | 1.040<br>(0.18)   | 1.025<br>(0.11)   |                                       | 1.040<br>(0.17)   | 1.025<br>(0.11)   |
| Visits children regularly                                                                  |                            |                   | 1.028<br>(0.12)   | 0.961<br>(0.17)   |                                       | 1.028<br>(0.13)   | 0.961<br>(0.17)   |
| Does not visit children regularly                                                          |                            |                   | 0.452<br>(2.86)** | 0.495<br>(2.61)** |                                       | 0.452<br>(2.47)*  | 0.495<br>(2.10)*  |
| Education: some college or more                                                            |                            |                   | 0.449<br>(3.35)** | 0.510<br>(2.59)** |                                       | 0.449<br>(3.51)** | 0.510<br>(2.81)** |
| Victim is wife - former and current                                                        |                            |                   | 0.704<br>(2.30)*  | 0.869<br>(0.73)   |                                       | 0.704<br>(1.90)   | 0.869<br>(0.72)   |
| Non-English Speaker                                                                        |                            |                   | 0.617<br>(2.80)** | 0.935<br>(0.31)   |                                       | 0.617<br>(2.02)*  | 0.935<br>(0.24)   |
| <i>Employment Status (employed as base comparison group)</i>                               |                            |                   |                   |                   |                                       |                   |                   |
| Employed part-time                                                                         |                            |                   |                   | 1.079<br>(0.32)   |                                       |                   | 1.079<br>(0.30)   |
| Not employed                                                                               |                            |                   |                   | 0.981<br>(0.11)   |                                       |                   | 0.981<br>(0.10)   |
| <i>Race/ethnicity (African American as base comparison group)</i>                          |                            |                   |                   |                   |                                       |                   |                   |
| Hispanic                                                                                   |                            |                   |                   | 0.553<br>(2.78)** |                                       |                   | 0.553<br>(2.36)*  |
| Other                                                                                      |                            |                   |                   | 0.460<br>(1.85)   |                                       |                   | 0.460<br>(1.84)   |
| White                                                                                      |                            |                   |                   | 0.689<br>(1.59)   |                                       |                   | 0.689<br>(1.36)   |
| <i>Total prior arrests of any offense (single, first offense as base comparison group)</i> |                            |                   |                   |                   |                                       |                   |                   |
| 2-3                                                                                        |                            |                   |                   | 2.647<br>(2.76)** |                                       |                   | 2.647<br>(2.49)*  |
| 4-5                                                                                        |                            |                   |                   | 2.194<br>(1.68)   |                                       |                   | 2.194<br>(1.75)   |
| >=6                                                                                        |                            |                   |                   | 3.700<br>(2.30)*  |                                       |                   | 3.700<br>(2.85)** |
| <i>Total prior arrests of DV offenses (single, first offense as base comparison group)</i> |                            |                   |                   |                   |                                       |                   |                   |
| 2                                                                                          |                            |                   |                   | 1.526<br>(1.79)   |                                       |                   | 1.526<br>(1.91)   |
| >=3                                                                                        |                            |                   |                   | 1.418<br>(1.28)   |                                       |                   | 1.418<br>(1.49)   |
| Had prior felony arrests                                                                   |                            |                   |                   | 1.013<br>(0.04)   |                                       |                   | 1.013<br>(0.04)   |
| Had prior arrests for drug offenses                                                        |                            |                   |                   | 0.802<br>(1.26)   |                                       |                   | 0.802<br>(1.08)   |
| Age at Intake                                                                              |                            |                   |                   | 0.961<br>(2.50)*  |                                       |                   | 0.961<br>(2.89)** |
| Age at First Arrest                                                                        |                            |                   |                   | 1.011<br>(0.54)   |                                       |                   | 1.011<br>(0.59)   |
| <i>CAGE Score (0 as base comparison group)</i>                                             |                            |                   |                   |                   |                                       |                   |                   |
| CAGE = 1                                                                                   |                            |                   |                   | 1.234<br>(0.82)   |                                       |                   | 1.234<br>(0.83)   |
| CAGE = 2                                                                                   |                            |                   |                   | 1.511<br>(1.61)   |                                       |                   | 1.511<br>(1.68)   |
| CAGE = 3                                                                                   |                            |                   |                   | 1.152<br>(0.55)   |                                       |                   | 1.152<br>(0.51)   |
| CAGE = 4                                                                                   |                            |                   |                   | 1.297<br>(1.10)   |                                       |                   | 1.297<br>(0.89)   |
| Program level variance                                                                     |                            |                   |                   |                   | 0.007<br>(0.24)                       | 0.009<br>(0.27)   | 0.016<br>(0.36)   |
| Observations                                                                               | 941                        | 941               | 941               | 941               | 941                                   | 941               | 941               |

Absolute value of z statistics in parentheses; \* significant at 5%; \*\* significant at 1%.

<sup>†</sup> In models 2 to 4, Huber-White sandwich estimate of variance is used for robust estimates of standard errors that adjust for intraclass correlations at the level of batterer intervention programs.

## Cognitive Behavioral and Supportive Group Treatments for Partner-Violent Men

TANYA M. MORREL

JEFFREY D. ELLIOTT

CHRISTOPHER M. MURPHY

CASEY T. TAFT

*University of Maryland, Baltimore County*

The study examined the relative efficacy of cognitive-behavioral group therapy (CBT) and supportive group therapy (ST) for partner-violent men at a community agency. Eighty-six men were assigned and exposed to ST or CBT. Outcome analyses, based on participant reports at pre- and posttreatment, collateral partner reports at pre, post, and 6-month follow-up, and criminal justice data gathered 2 to 3 years after treatment, revealed no significant differences between ST and CBT on the primary outcomes of partner aggression and arrests. Across conditions, clients showed significant decreases in physical assault, psychological aggression, and injuries, significant increases in self-esteem and self-efficacy for abstaining from partner aggression, and significant movement on stage-of-change scales. ST clients had significantly greater increases than CBT clients on two secondary outcome variables: negotiation tactics and self-efficacy for abstaining from verbal aggression. Neither partner reports of criminal recidivism nor criminal data revealed significant treatment condition differences. The findings failed to demonstrate an added benefit of behavioral group interventions over the effects of a supportive group treatment experience for partner-violent men.

Adult intimate partner violence has become an important clinical concern. Although more than 1,000 counseling programs for partner violence perpetrators have developed in the U.S., few outcome studies have examined the efficacy of such interventions. Early, uncontrolled studies reported favorable effects of treatment for domestic abuse perpetrators (Faulkner, Stoltenberg,

This research was supported by grants from the National Institutes of Health (R03-MH56373) and the Designated Research Initiative Fund of the University of Maryland, Baltimore County. The authors wish to thank Dr. Jerry Jennings for training project therapists in supportive therapy, and the staff of the Domestic Violence Center of Howard County, MD, for their support of our clinical research efforts.

Address correspondence to Christopher Murphy, Department of Psychology, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250; e-mail: chnmurphy@umbc.edu.

Cogen, Nolder, & Shooter, 1992; Hamberger & Hastings, 1988; Saunders & Hanusa, 1986). Studies with no or minimal treatment control groups, however, have yielded mixed results. Some have found that abuser treatment significantly reduces violence recidivism as assessed by victim report or criminal justice data (Dutton, 1986; Palmer, Brown, & Barrera, 1992), whereas others have found no significant added benefit from treatment relative to intensive monitoring or no-treatment controls (e.g., Dunford, 2000). Furthermore, violence recidivism rates are high during the year after treatment, usually between 25% and 50% by victim report.

Most of the available treatment models contain elements of cognitive and behavioral therapies (CBT). Often included are anger reduction techniques involving relaxation and cognitive restructuring and communication skills or assertiveness training. CBT interventions for intimate partner violence are often couched within a feminist framework that views abusive behavior as men's control of women in the domestic sphere (Saunders, 1996).

Although less common in the field, unstructured, supportive group therapies (ST) have also been developed for partner-violent men. These programs encourage the abusive client to develop a personal agenda for change. The therapists facilitate mutual support and encourage peer remediation of abusive behavior and relationship problems. The goal is to create a natural setting for clients to improve their communication, express feelings appropriately, reduce isolation and excessive dependence on the partner, and alleviate humiliation and shame in relationships with other men (Jennings, 1987; Jennings & Murphy, 2000). In a recent client satisfaction survey at abuser counseling programs, the most frequent recommendation was to create a more supportive treatment atmosphere (Gondolf & White, 2000).

For the current study, this relatively unstructured approach provided an alternative treatment that could be contrasted with the theoretically active elements of CBT. To date, studies of abuser intervention have not adequately examined the specific effects of CBT interventions. In particular, there have been important confounds in studies that have compared group CBT to control conditions that offer the benefits of participating in a group (e.g., receiving social support, reducing isolation, discussing problems) but do not provide specific skills training.

One relevant study compared a structured psychoeducational group delivered by professional therapists to a self-help group run by a former abusive client or a combination of these two treatments (Edleson & Syers, 1990). No significant differences emerged between conditions in the intensity or frequency of violence at follow-up as reported by abusive clients and their partners. At the 6-month follow-up, men in the structured psychoeducational group made significantly fewer terrorizing threats than men in the self-help or combined groups. However, close to half of participants failed to complete at least 80% of sessions. Outcome data were available on less than half of the sample at each follow-up and consisted of self-reports of abuse for some cases and victim reports for other cases. In addition, therapist professional

training was confounded with treatment condition and no treatment integrity data were provided. The impact of specific CBT interventions was therefore unclear in this study.

A second study compared a feminist CBT group program to a process-psychodynamic group approach (Saunders, 1996). No significant differences were found in recidivist violence at follow-up (averaging 2 years posttreatment), as assessed by victim report supplemented with self-report and criminal justice data. Interestingly, CBT was more effective than psychodynamic treatment for clients with antisocial personality features, whereas the psychodynamic groups were more effective for clients with dependent personality features. Unfortunately, the findings may have been influenced by attrition from treatment, which was high overall (about 40%) and significantly lower in the process-psychodynamic condition. In addition, the psychodynamic group included structured homework assignments and active therapist interpretations of problem behaviors, limiting its value as a control for the active skills-training features of CBT.

The current study was designed to determine whether a structured, skills training group based on the principles of CBT was more effective than unstructured, supportive group therapy in reducing rates of physical and psychological abuse and in affecting secondary treatment targets that may confer risk for continued problems with abuse, namely, self-esteem, self-efficacy, readiness to change abusive behavior, and communication difficulties. Efforts were made to limit treatment dropout, examine treatment adherence, control for therapist effects, and obtain outcome data from multiple sources.

Both the CBT and ST approaches were conducted by a male-female co-therapy team in a close-ended group format comprising 16 weekly 2-hour sessions. The CBT approach (Murphy & Scott, 1995) addressed, in sequence: (a) the motivation to end abusive behavior and commitment to nonviolent relationships; (b) crisis-management strategies such as time-out; (c) anger-management techniques including self-monitoring of anger and coping, relaxation training, and cognitive restructuring of angry thoughts; and (d) communication skills, including active listening, emotional expression, assertiveness, and compromise. Time was allotted during each session to discuss ongoing personal and relationship issues not addressed during the structured portion of the session. The program is based on social learning theory (Bandura, 1986). It incorporated elements of other CBT and motivational treatments (Miller & Rollnick, 1991; Novaco, 1975; Rosenbaum & O'Leary, 1986; Sonkin & Durphy, 1989) and was informed by research on the individual and dyadic correlates of relationship abuse (e.g., Holtzworth-Munroe & Hutchinson, 1993; Margolin, John, & Gleberman, 1988).

The ST approach was based on the work of Jennings (1987) and on Yalom's (1995) primary therapeutic factors for group treatment. ST therapists provided brief training in the crisis-management skill of time-out during an early treatment session, but otherwise provided minimal therapist-directed intervention beyond encouragement of a mutually supportive environment

and focus on relationship issues and ending abusive behavior. The therapists allowed group members to set the session agenda and addressed themes and topics that emerged spontaneously in the group interaction. The therapists emphasized a collaborative group norm and refrained from using active skill-training interventions. Therapists were instructed to address the group as a whole rather than individuals, and to use brief verbalizations and nonverbal gestures to stimulate vigorous and helpful group interactions.

Because of the limited prior data suggesting that either therapy would be more effective than the other, the study was largely exploratory. The working hypothesis was that the active elements of CBT would produce greater decreases in psychological and physical aggression and greater increases in self-efficacy for abstaining from aggressive behavior in difficult relationship situations when compared to ST. The investigators' primary allegiance was to the CBT model, but an expert in ST for abusive men (Dr. Jerry Jennings) trained project therapists in the delivery of this treatment.

## Method

### *Participants*

Participants were 86 men seeking group treatment for partner-abusive behavior at the Domestic Violence Center (DVC) of Howard County, Maryland. Participants met the following inclusion criteria: (a) age of 18 or over; (b) displayed no signs of psychosis during intake (i.e., no hallucinations, delusions, or formal thought disorder); (c) had a documented problem with relationship abuse, as indicated by a history of physical relationship aggression reported at baseline by self or partner, or an arrest report with clear documentation of partner violence; and (d) provided written consent for study participation and collateral partner contact. Collateral partners provided research assent during an initial phone interview. The treatment agency maintained additional requirements for enrollment in treatment: (1) If the participant reported current use of psychotropic medication, permission for group referral was secured from the treating psychiatrist and/or primary therapist; (2) if the participant reported use of stimulants, PCP, or cocaine in the past 6 months, he was required to be enrolled in substance abuse treatment before referral to group; and (3) the participant received an alcohol abuse evaluation and complied with resulting recommendations if (a) either he or his partner indicated an active problem with alcohol, (b) he scored greater than 10 on the Alcohol Use Disorders Identification Test (AUDIT; Babor, de la Fuente, Saunders, & Grant, 1992), or (c) the participant or his partner reported more than one episode of domestic violence while the participant was under the influence of alcohol.

During the recruitment period, 118 men presented for intake. Two refused to participate in research, 20 failed to complete the intake process or refused treatment during the intake process, 5 were not assigned to group because of schedule conflicts or failure to comply with a referral for drug and alcohol

treatment, and 5 were assigned to group but did not attend any group sessions. A total of 86 men (38 in ST and 48 in CBT) attended at least one group session and constituted the “exposed to treatment” sample. Of these, 74 (33 in ST and 41 in CBT) received a credible dose of treatment (12 or more sessions). Results from the credible dose of treatment sample are reported with footnotes indicating any differences in findings for the exposed-to-treatment sample.<sup>1</sup>

Participants’ average age was 34.7 years ( $SD = 7.8$ ). They had an average of 13.1 years of formal education ( $SD = 2.7$ ), and net monthly income of \$1,800 ( $SD = \$1,900$ ). A court recommendation to treatment was in force for 74% of cases; 7% had a court case pending at the time of intake, and 19% had no court involvement. With regard to ethnicity, 60% of the sample were non-Hispanic Caucasian, 30% African American, 3% Asian American, 2% Native American, 1% Hispanic, and 2% “other.”

## Procedure

### *Assignment to Conditions*

After the intake, participants were assigned to the next scheduled group. Assignment to condition was unsystematic, but not technically random. CBT and ST groups were alternated, with deviations from alternation arranged in order to counterbalance therapists across study conditions. Possible selection biases were controlled by assigning treatment groups to CBT or ST prior to constituting the group membership. Group size varied from 7 to 10 clients based on fluctuations in client flow into regularly scheduled groups. A total of 5 ST and 5 CBT groups were conducted.

### *Therapists and Treatment Integrity*

Each group was co-led by a clinical psychology graduate student trainee and one of two agency staff who had master’s-level training and prior supervised experience administering the CBT protocol. All group therapists read the treatment manuals and received formal classroom training in general principles of treatment for partner-violent men (5 hours), CBT (8 hours), and ST (8 hours). Co-therapist teams met weekly with the supervising psychologist to review session videotape segments, discuss difficulties in administering the protocols, monitor treatment adherence, and receive guidance for upcoming sessions. The two master’s-level therapists were counterbalanced across conditions.

<sup>1</sup> Some study participants reported abuse toward more than one victim, or reported being in a new relationship with someone other than the original identified victim. In these cases, an attempt was made to contact and complete an assessment at each time point with each relationship partner. Given that detection of problem behavior was the main priority, the analyses used the data from whichever partner reported the greatest frequency of physical assault on the CTS2 at each assessment, or the highest level of psychological aggression if no physical assault was reported. Four pretreatment collateral informants were substituted with different informants at posttreatment, and no further substitutions were needed at 6-month follow-up.

Treatment adherence was evaluated from key therapist activities expected to be more common in either CBT or ST. For each 16-session group, one session was randomly sampled from weeks 1 to 4, 5 to 8, 9 to 12, and 13 to 16. Sessions were divided into forty 3-minute segments. Raters (undergraduate assistants who were blind to the nature of the study and treatment conditions) coded a therapist activity as present if either therapist engaged in it during the 3-minute segment. Raters could code multiple therapist behaviors within each segment, yielding a continuous measure on a scale from 0 to 40 for each therapist activity at the session level. Raters were trained to 90% agreement with a set of criterion tapes on specific therapist behaviors. Interrater agreement, assessed on 25% of the tapes (one from each treatment group, selected randomly), revealed 89% overall agreement and an overall average category kappa of .45.

Adherence was analyzed by discriminant function of the session-level data, with two composite therapist activity variables as the predictors, and treatment condition as the criterion. The composite variables were the total number of therapist activities specified, a priori, to be representative of CBT (e.g., leads role-play, reviews homework, assigns homework, explores details, makes presentation), and ST (e.g., facilitates individual expression, encourages exploration, comments on interactions, elicits reactions or feelings from the group, draws parallels between group members). Adherence analyses examined both the frequency and relative proportion of total therapist behaviors coded for the ST and CBT behavior categories.

The discriminant analysis of therapist proportion variables was significant (chi-square = 36.95,  $df = 2$ ,  $N = 40$ ,  $p < .001$ ). In ST, therapists allocated only 1/20 of their coded behavior to the techniques prescribed for CBT (mean = .05,  $SD = .09$ ), whereas in CBT therapists allocated 1/3 of their coded behavior to these directive intervention techniques (mean = .34,  $SD = .13$ ). Results of the second analysis examining the frequency of therapist behaviors were also significant (chi-square = 42.09,  $df = 2$ ,  $N = 40$ ,  $p < .001$ ). In CBT, therapists were more active, engaging in considerably more coded behaviors overall. Interestingly, the therapist behaviors specified, a priori, to reflect ST, although representing a greater proportion of overall therapist behavior in the ST condition, actually occurred at roughly similar frequency rates in both treatments (in ST: mean = 42.2,  $SD = 17.3$ ; in CBT: mean = 50.9,  $SD = 19.3$ ). In contrast, therapist behaviors specified, a priori, to reflect CBT occurred almost exclusively in the CBT condition (in ST: mean = 2.4,  $SD = 5.0$ ; in CBT: mean = 26.3,  $SD = 10.8$ ). Thus, ST appears to have provided a good control condition for the common factors of supportive group therapist interventions. Note, however, that therapists were also trained to use nonverbal techniques in ST designed to facilitate and stimulate therapeutic support among group members. Examples include regular eye contact with all group members, nonverbal gestures to refer client questions to the group, inviting looks to elicit comments from specific group members, or changes in posture and expression to display enthusiasm for meaningful group transactions. These were too subtle to include in the adherence coding

system. Nevertheless, the adherence data are consistent with the clinical observation that the ST groups engaged in lengthier client-directed interaction sequences involving more limited therapist verbalizations, and CBT-consistent therapist behaviors were extremely rare in the ST condition.

### *Measures*

*Standardized measure of aggression.* The Revised Conflict Tactics Scale (CTS2; Straus, Hamby, Boney-McCoy, & Sugarman, 1996) was used to assess partner aggression. The CTS2 has five subscales: Negotiation (6 items), Sexual Coercion (7 items), Injury (6 items), Psychological Aggression (8 items), and Physical Assault (12 items). The CTS2 scales have good internal consistency, and evidence of their concurrent, content, and construct validity has been presented (Straus et al., 1996). Participants and collaterals indicated the frequency with which the participant had engaged in these behaviors on a scale from *never* to *more than 20 times* during the 6 months prior to intake (baseline), the approximately 6-month period from intake to the end of treatment (posttreatment), and the 6 months after treatment (follow-up).

*Criminal recidivism* was assessed through review of participants' criminal histories, obtained 22 to 36 months after the scheduled completion of treatment from an electronic database for the state of Maryland. Criminal histories were obtained for 63 of the 86 participants in the exposed-to-treatment sample.<sup>2</sup> Using a method outlined in prior research (Murphy, Musser, & Maton, 1998), criminal recidivism was coded on the basis of charges that are often received as a result of domestic abuse incidents, although specific information on the relationship to the victim of the crime was not available in the electronic database. Most of these charges are violent or threatening in nature.

*Incidence measures of aggression.* Partners answered structured interview questions at each assessment about the number of incidents of physical abuse, arrests, and police visits during the subsequent assessment interval.

*Global impressions of change.* A scale to indicate global partner impressions of treatment outcome was created by standardizing and summing 7 partner interview items assessed at posttreatment and follow-up. Two items inquired whether the frequency and intensity of abuse had decreased, increased, or stayed about the same. Five items inquired about changes for the better or worse on a 5-point scale for level of physical aggression, verbal aggression, fear, general treatment by the client, and relationship quality. The resulting global index had good internal consistency (coefficient alpha = .89 at posttreatment and .90 at follow-up). The validity of this global index of change is indicated by significant correlations with partner reports at posttreatment on CTS2 scales for physical assault, injuries, and psychological aggression, and on the verbal problem checklist (*r*s ranging from .34 to .58).

<sup>2</sup> Criminal justice data could not be located for several reasons, including an absence of criminal history for some cases who were self-referred or involved in civil court proceedings, inadequate tracking information, or criminal history in a state other than Maryland.



At the 6-month follow-up, the global index was correlated significantly with the verbal problem checklist ( $r = .43$ ), but not with the CTS2 scales ( $r$ s ranging from  $-.10$  to  $.17$ ). The global index of outcome was reasonably stable from posttreatment to 6-month follow-up ( $r = .51$ ).

*Communication behaviors.* Collateral partner reports on the Verbal Problem Checklist (VPC; Haynes, Chavez, & Samuel, 1984) were used to assess the client's communication difficulties. Collateral reports on the VPC have been shown to correlate highly with observer ratings of negative communication during a problem-solving task. The current study used a 15-item subscale identified by Vivian and Langhinrichsen-Rohling (1994) that had high internal consistency ( $\alpha = .91$ ) and a homogeneous item pool reflecting negative communication behaviors.

The 6-item CTS2 Negotiation subscale was used to assess the client's use of negotiation skills in response to conflict, as assessed by self and partner report. This scale is a face valid measure of nonaggressive conflict behaviors that is relatively independent of partner aggression (Straus et al., 1996).

*Readiness to change.* The Safe At Home Instrument (Begun et al., in press), a 35-item self-report questionnaire, was used to assess men's readiness to change intimate partner violence at pretreatment and posttreatment. Items were developed to reflect the stages of change (Prochaska, DiClemente, & Norcross, 1992). Respondents rate agreement with each item on a 5-point scale. Data from a large sample of treatment-seeking partner-violent men revealed three factors, with evidence of factor consistency and measurement invariance across treatment sites and multiple administrations. The subscales assess precontemplation (e.g., "It's no big deal if I lose my temper from time to time"); contemplation (e.g., "I want to do something about my problem with conflict"); and preparation/action (e.g., "Even though I get angry, I know ways to keep from losing control"). Internal consistency in the development sample was .91, .79, and .59 for the Contemplation, Preparation/Action, and Precontemplation scales, respectively. A readiness-to-change composite is calculated by adding the contemplation and preparation/action scores and subtracting the precontemplation score, using the seven highest loading items for each subscale. Validity was examined using data from the current sample (Begun et al., in press). Example findings include a negative correlation between Precontemplation scale scores and assumption of responsibility for abusive behavior, a positive correlation between Preparation/Action scores and self-efficacy for abstaining from verbal aggression, and higher Contemplation scale scores among self- versus court-referred clients.<sup>3</sup>

*Self-esteem* was measured by self-report at pre- and posttreatment using the 10-item Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965). Items were administered on a 4-point Likert scale, ranging from *strongly agree* (1)

<sup>3</sup> This previous work examined convergent, discriminant, and predictive validity for the safe-at-home measure in the current sample, whereas the current study examines the measure as a treatment outcome variable.

to *strongly disagree* (4). The RSES is a reliable and valid indicator of global self-esteem (Demo, 1985; Fleming & Courtney, 1984). Convergent validity has been demonstrated via significant correlations with other self-report measures (e.g., Coopersmith Self-Esteem Inventory) and peer ratings of self-esteem.

*Self-efficacy* at pre- and posttreatment was assessed with a 15-item questionnaire. Participants rated on a 7-point Likert scale how confident they were that they could handle difficult relationship situations without becoming aggressive (*not at all confident* to *extremely confident*). Separate ratings were made for verbal and physical aggression. In posttreatment data from 35 abusive men, an earlier, 23-item version of the scale had high internal consistency. Self-efficacy was negatively correlated with the tendency to hold anger in, but had near zero correlations with reports of physical and verbal aggression during the preceding months. Perpetrator reports of posttreatment self-efficacy were negatively correlated with victim reports of subsequent abuse in the follow-up period ( $r = -.19$  and  $-.39$  for verbal and physical aggression, respectively; Morrel & Murphy, 1996). In the current sample, the 15-item scale had high internal consistency (coefficient alpha = .93 and .95 for verbal and physical aggression, respectively).

## Results

### *Characteristics of the Recruited Sample*

Those who presented for program intake but did not become part of the study did not differ significantly from the study participants in age ( $t = .02$ ,  $df = 114$ ,  $p = .98$ ), years of education ( $t = .71$ ,  $df = 111$ ,  $p = .48$ ), income ( $t = 1.05$ ,  $df = 111$ ,  $p = .30$ ), ethnic minority status (white vs. nonwhite: chi-square = .99,  $N = 116$ ,  $df = 1$ ,  $p = .32$ ), court-mandated referral status (mandated vs. case pending vs. nonmandated: chi-square = 2.8,  $N = 116$ ,  $df = 2$ ,  $p = .24$ ), or pretreatment partner reports of physical assault ( $t = .21$ ,  $df = 91$ ,  $p = .84$ ) and psychological aggression ( $t = .61$ ,  $df = 92$ ,  $p = .54$ ).

### *Group Assignment*

The assignment to conditions appears to have achieved comparability with respect to demographics and initial problem levels. There were no significant differences between ST and CBT participants in age ( $t = .48$ ,  $df = 85$ ,  $p = .64$ ), education ( $t = .92$ ,  $df = 82$ ,  $p = .36$ ), income ( $t = .82$ ,  $df = 83$ ,  $p = .42$ ), ethnic minority status (chi-square = .19,  $N = 86$ ,  $df = 1$ ,  $p = .66$ ), court-mandated referral status (chi-square = 2.6,  $N = 86$ ,  $df = 2$ ,  $p = .27$ ), or pretreatment partner reports of physical assault ( $t = .34$ ,  $df = 76$ ,  $p = .74$ ) and psychological aggression ( $t = .27$ ,  $df = 76$ ,  $p = .79$ ).

### *Sample Attrition*

The proportion of the total sample attending at least 75% of scheduled sessions was high for this population (86%). There were no significant differences

between treatment conditions in the proportion of clients who received a credible dose of treatment (chi-square = .04,  $N = 88$ ,  $df = 1$ ,  $p = .85$ ). Men who dropped out of therapy (i.e., completed fewer than 12 of 16 sessions) did not differ significantly from treatment completers in minority status (white vs. nonwhite; chi-square = 1.23,  $N = 86$ ,  $df = 1$ ,  $p = .27$ ), education ( $t = 1.47$ ,  $df = 83$ ,  $p = .13$ ), income ( $t = 1.52$ ,  $df = 84$ ,  $p = .20$ ), or pretreatment levels of physical assault ( $t = 1.48$ ,  $df = 77$ ,  $p = .14$ ) and psychological aggression ( $t = 0.49$ ,  $df = 77$ ,  $p = .62$ ). However, men who dropped out of treatment were significantly younger (mean = 30.48;  $SD = 6.63$ ) than those who completed treatment (mean = 35.40,  $SD = 7.79$ ;  $t = 2.12$ ,  $df = 86$ ,  $p < .05$ ), and less likely to be court mandated (58% vs. 77%; chi-square = 7.04,  $N = 88$ ,  $df = 2$ ,  $p < .05$ ). There were no significant differences between ST and CBT conditions in the proportion of partners assessed at baseline (chi-square = 0.91,  $N = 88$ ,  $df = 1$ ,  $p = .34$ ), posttreatment (chi-square = 0.02,  $N = 88$ ,  $df = 1$ ,  $p = .98$ ), or 6-month follow-up (chi-square = 1.12,  $N = 88$ ,  $df = 1$ ,  $p = .29$ ). Overall, partners were successfully interviewed for 91% of cases at baseline, 72% of cases at posttreatment, and 61% of cases at follow-up.

#### *Tests of Changes Over Time Using Partner Reports of Abuse*

Change over time across treatment conditions in partner report of physical aggression, psychological aggression, injuries, sexual coercion, negotiation skills, and verbal problems was examined using repeated measure ANOVAs. Log transformations of the physical assault, injury, and sexual coercion variables were used to correct for nonnormality. Means and standard deviations of the nontransformed variables are reported in Table 1. Across the two treatment conditions, physical assault, psychological aggression, injuries, and sexual coercion decreased significantly from pre- to posttreatment,  $F(1, 52) = 48.78$ ,  $p < .001$ ,  $F(1, 50) = 12.29$ ,  $p < .01$ ,  $F(1, 52) = 34.76$ ,  $p < .001$ ,  $F(1, 52) = 7.23$ ,  $p < .05$ , respectively, and from pretreatment to 6-month follow-up,  $F(1, 44) = 56.74$ ,  $p < .001$ ,  $F(1, 36) = 15.06$ ,  $p < .001$ ,  $F(1, 44) = 40.59$ ,  $p < .001$ ,  $F(1, 44) = 9.24$ ,  $p < .01$ , respectively. Across conditions, there were nonsignificant changes in perpetrators' levels of negotiation behaviors from pre to post,  $F(1, 50) = .02$ ,  $p = .97$ , and from pre to follow-up,  $F(1, 36) = 1.50$ ,  $p = .23$ ; and verbal problems from pre to post,  $F(1, 49) = 3.39$ ,  $p = .07$ , and from pre to follow-up,  $F(1, 35) = .29$ ,  $p = .60$ .

#### *Tests of Treatment Condition Differences Using Partner Reports of Abuse*

Differences between treatment conditions for these measures were examined using ANCOVAs with the pretreatment level of the variable as the covariate. The only significant difference between conditions was in partner reports of the use of negotiation skills by perpetrators,  $F(1, 48) = 5.00$ ,  $p < .05$ , which was greater in ST than in CBT at posttreatment (see Table 1).<sup>4</sup>

<sup>4</sup> In the exposed-to-treatment sample this difference was not significant, although the trend was the same.

TABLE 1  
MEANS AND STANDARD DEVIATIONS FOR PARTNER- AND SELF-REPORTS OF OUTCOME BY TREATMENT CONDITION

| Variable                                       | ST<br>M (SD)  |               |               | CBT<br>M (SD) |               |               |
|------------------------------------------------|---------------|---------------|---------------|---------------|---------------|---------------|
|                                                | Pretreatment  | Posttreatment | Follow-up     | Pretreatment  | Posttreatment | Follow-up     |
| Partner report                                 |               |               |               |               |               |               |
| Physical assault (CTS2) <sup>†</sup>           | 9.86 (13.07)  | 1.12 (2.55)   | .56 (2.12)    | 10.45 (14.85) | .90 (2.21)    | .95 (2.48)    |
| Psychological aggression (CTS2)                | 18.80 (13.39) | 9.79 (9.95)   | 10.28 (10.28) | 16.88 (11.94) | 12.31 (8.76)  | 11.95 (6.82)  |
| Injury (CTS2) <sup>†</sup>                     | 3.18 (4.78)   | .40 (1.12)    | .17 (.71)     | 3.80 (5.07)   | .42 (1.20)    | .24 (1.09)    |
| Sexual coercion (CTS2) <sup>†</sup>            | .91 (2.36)    | .57 (1.53)    | .00 (.00)     | 3.25 (6.51)   | .39 (1.28)    | .57 (1.80)    |
| Negotiation (CTS2)                             | 16.77 (9.14)  | 20.96 (9.68)  | 14.39 (8.91)  | 17.26 (9.77)  | 16.48 (9.97)  | 15.90 (10.63) |
| Verbal problems (VPC)                          | 42.20 (15.50) | 33.28 (19.05) | 42.41 (16.62) | 42.36 (15.50) | 40.39 (15.31) | 41.48 (14.17) |
| Self-report                                    |               |               |               |               |               |               |
| Physical assault (CTS2) <sup>†</sup>           | 2.83 (5.71)   | .19 (.64)     | —             | 3.93 (6.15)   | 1.40 (5.18)   | —             |
| Psychological aggression (CTS2)                | 9.75 (7.97)   | 3.34 (3.84)   | —             | 9.39 (7.83)   | 4.58 (6.50)   | —             |
| Injury (CTS2) <sup>†</sup>                     | 1.03 (1.99)   | .19 (.91)     | —             | 1.53 (2.37)   | .80 (2.97)    | —             |
| Sexual coercion (CTS2) <sup>†</sup>            | .77 (2.09)    | .23 (1.09)    | —             | .85 (2.06)    | .98 (2.92)    | —             |
| Negotiation (CTS2)                             | 16.45 (8.58)  | 16.01 (7.36)  | —             | 18.38 (9.99)  | 19.15 (10.78) | —             |
| Physical aggression self-efficacy <sup>†</sup> | 78.67 (27.91) | 94.53 (19.23) | —             | 87.12 (23.75) | 92.58 (24.16) | —             |
| Verbal aggression self-efficacy                | 59.77 (27.19) | 87.30 (15.88) | —             | 71.44 (24.82) | 81.22 (25.60) | —             |
| Self-esteem                                    | 33.87 (4.53)  | 35.25 (3.99)  | —             | 32.50 (5.05)  | 34.07 (4.56)  | —             |
| Readiness to change                            | 4.43 (1.11)   | 4.69 (1.75)   | —             | 4.54 (1.49)   | 4.81 (1.66)   | —             |
| Stages of change: precontemplation             | 2.72 (.47)    | 2.49 (.66)    | —             | 2.62 (.59)    | 2.56 (.50)    | —             |
| Stages of change: contemplation                | 3.50 (.81)    | 3.23 (.89)    | —             | 3.61 (.77)    | 3.36 (.97)    | —             |
| Stages of change: action/preparation           | 3.78 (.74)    | 4.01 (.72)    | —             | 3.68 (.67)    | 4.08 (.58)    | —             |

Note. CTS2 = Revised Conflict Tactics Scale; VPC = Verbal Problem Checklist.

<sup>†</sup> Log normal transformation of variable used in analyses due to nonnormality but means of nontransformed variables reported.

To assess differences between conditions in incidence measures of aggression at posttreatment and follow-up, ANCOVAs were conducted with pretreatment measures entered as covariates. All incidence measures were log transformed to correct for nonnormality. Means and standard deviations of the nontransformed variables are reported in Table 2. Partner reports of police visits, arrests, and physical abuse incidents did not differ between conditions at posttreatment or follow-up.

On the global index of partner outcome impressions (assessing perceived change from before treatment), simple ANOVAs revealed no significant differences between CBT and ST at posttreatment or follow-up (see Table 2). Item responses at posttreatment revealed that across conditions 76% of partners reported that the perpetrator was engaging in less physical aggression (8% "somewhat less"; 67% "much less"); 64% reported less verbal aggression (19% "somewhat less"; 45% "much less"); 71% reported less fear (19% "somewhat less"; 54% "much less"); 56% reported improvement in how the partner was treating them (20% "somewhat better"; 36% "much better"); and 63% reported improvement in the overall quality of the relationship (25% "somewhat better"; 37% "much better").

Finally, there were no differences between CBT and ST in clinically significant change, assessed as no violence outcomes. At posttreatment, 74% of partners in both conditions reported no physical violence on the CTS2 (chi-square = 0.00,  $N = 62$ ,  $p = .98$ ). During the 6 months after treatment, 90% of partners in the ST condition and 79% of partners in the CBT condition reported no physical violence on the CTS2 (chi-square = 1.11,  $N = 50$ ,  $p = .29$ ).

TABLE 2  
MEANS AND STANDARD DEVIATIONS FOR INCIDENCE MEASURES AND GLOBAL IMPRESSIONS  
OF CHANGE BY TREATMENT CONDITION

| Variable                                                 | Posttreatment      |                    | Follow-up          |                    |
|----------------------------------------------------------|--------------------|--------------------|--------------------|--------------------|
|                                                          | ST                 | CBT                | ST                 | CBT                |
|                                                          | $M^{\dagger}$ (SD) | $M^{\dagger}$ (SD) | $M^{\dagger}$ (SD) | $M^{\dagger}$ (SD) |
| Number of physical abuse incidents <sup>†</sup>          | .67 (1.81)         | .83 (2.17)         | .53 (1.37)         | 6.14 (26.12)       |
| Number of police visits <sup>†</sup>                     | .16 (.62)          | .36 (.62)          | .24 (.75)          | .29 (.56)          |
| Number of arrests (partner report) <sup>†</sup>          | .12 (.60)          | .10 (.31)          | .00 (.00)          | .00 (.00)          |
| Global Index of Outcome                                  | .06 (.75)          | -.10 (.70)         | .13 (.86)          | -.09 (.76)         |
| Number of new charges from criminal records <sup>†</sup> |                    |                    | .10 (.30)          | .13 (.43)          |

*Note.* All variables are from collateral partner report, with the exception of new charges from criminal records. Partner follow-up was conducted 6 months after treatment; criminal records cover a period that ranged from 24 to 36 months after treatment. For partner reports at posttreatment and follow-up,  $N = 52$  and 36, respectively. For criminal data  $N = 63$ .

<sup>†</sup> This variable underwent log transformation prior to analysis. The table presents data on the untransformed version.

*Tests of Treatment Condition Differences Using Criminal Reports*

Recidivist incidents were present during follow-up for 11 of 63 cases (17%). They had a maximum of three arrest incidents during follow-up and had received the following range of charges: assault, battery, violation of a protection order, malicious destruction of property, child abuse, telephone abuse, and assault with a deadly weapon. The number of arrest incidents during the follow-up was analyzed, with lifetime arrests on similar charges in Maryland prior to treatment as a covariate. The participants had an average of 1.83 incidents prior to treatment ( $SD = 1.65$ ; range = 0 to 7). These variables were log transformed to reduce skew and kurtosis. Results revealed no significant difference between treatment conditions in the number of arrests during follow-up,  $F(1, 49) = .04, p = .85$ .

*Tests of Changes Over Time Using Client Self-Report Data*

Repeated measure ANOVAs were used to test for differences in change over time across treatment conditions on client self-reports of physical aggression, psychological aggression, injuries, sexual coercion, negotiation skills (CTS2), self-efficacy in abstaining from physical and verbal aggression, and self-esteem.<sup>5</sup> Physical assault, injury, sexual coercion, and confidence in abstaining from physical aggression were log transformed to correct for nonnormality. Means and standard deviations of the nontransformed variables are reported in Table 1. Significant reductions from pre- to posttreatment were found for client reports on the CTS2 scales for physical assault,  $F(1, 67) = 30.09, p < .001$ , psychological aggression,  $F(1, 66) = 49.06, p < .001$ , and injury,  $F(1, 67) = 13.86, p < .001$ . Sexual coercion rates did not decrease significantly from pre- to posttreatment, most likely due to the low rates reported by clients at pretreatment,  $F(1, 67) = 1.58, p = .21$ . Similarly, the change in self-reports of negotiation was not significant,  $F(1, 67) = .09, p = .77$ . Participants' self-esteem,  $F(1, 72) = 9.58, p < .01$ , and self-efficacy for abstaining from verbal aggression,  $F(1, 69) = 26.09, p < .001$ , increased significantly from pre- to posttreatment across conditions. However, the change in participants' self-efficacy for abstaining from physical aggression was not significant,  $F(1, 64) = 3.67, p = .06$ . Although there was no significant increase over time on the overall readiness-to-change composite,  $F(1, 65) = 1.40, p = .24$ , or the Precontemplation scale,  $F(1, 70) = 3.58, p = .06$ , there was a significant increase on the Preparation/Action scale,  $F(1, 69) = 12.86, p < .01$ , and a significant decrease on the Contemplation scale,  $F(1, 67) = 12.06, p < .01$ .

*Tests of Treatment Condition Differences Using Client Self-Report Data*

Differences between treatment conditions in self-reports were examined using ANCOVAs with the pretreatment level of the variable as the covariate.

<sup>5</sup> Because of resource constraints, the decision was made to assess male clients at baseline and posttreatment only, and to concentrate follow-up data collection on collateral partners. Therefore, analyses of client self-report data examine only pre- and posttreatment.

The only significant difference was in posttreatment levels of self-efficacy for abstaining from verbal aggression,  $F(1, 67) = 4.83, p < .05$ . Although the posttreatment levels of self-efficacy were only slightly higher in ST versus CBT (see Table 1), CBT participants reported higher self-efficacy at pretreatment,  $F(1, 68) = 3.17, p = .08$ . The posttreatment difference between ST and CBT was significant only after controlling for baseline levels, suggesting that the absolute level of self-efficacy at posttreatment was not different, but the degree of change over time was greater in ST.

## Discussion

Both CBT and ST were associated with significant reductions in physical assault, psychological aggression, injuries, and sexual coercion. These changes were observed at the end of the 16-session treatment and were maintained at 6-month follow-up. Counter to predictions, CBT was not more effective than ST in reducing abusive behavior. Similarly, with regard to secondary treatment targets such as self-esteem and self-efficacy for abstaining from physical aggression, significant changes were observed from pre- to posttreatment with no differences between conditions. At posttreatment, the majority of collateral partners across conditions provided global impressions of positive change.

The only exceptions to this general pattern seemed to favor the supportive therapy condition. At posttreatment, ST was associated with greater self-reported self-efficacy for abstaining from verbal aggression and higher collateral partner reports of clients' use of nonaggressive negotiation tactics to handle relationship conflict. It is important to note, however, that these variables were not considered primary outcomes, and their implications for abusive behavior are not entirely clear. Higher levels of negotiation, in particular, may reflect a greater degree of ongoing conflict or an enhanced propensity to handle conflict constructively. Given the large number of analyses comparing CBT and ST, it is not surprising that two significant differences emerged. The number and pattern of significant results are not sufficient to warrant firm conclusions about the superiority of either treatment.

Consistent with prior pre- to posttreatment studies of partner-violent men, reductions in both psychological and physical aggression over the course of treatment were relatively large in magnitude. Unfortunately, given the absence of a no-treatment control group, these changes cannot be attributed directly to the treatments, and may result from naturally occurring change processes, sample attrition, response to legal sanctions, maturation, or repeated testing. Studies of similar CBT group programs for abusive men containing no-treatment or minimal treatment controls have yielded mixed findings of efficacy (Dunford, 2000; Dutton, 1986; Palmer et al., 1992; Waldo, 1988), and a recent meta-analysis has reported average effects of small magnitude (Babcock & LaTaillade, 2000). An explanation involving natural fluctuations in the low-frequency behavior of partner violence is less consistent with the current study findings, as changes were found on higher frequency behaviors

such as psychological aggression and on relevant cognitive and attitudinal measures of the change process as well. Even though a specific added benefit of CBT interventions was not observed, the findings suggest that the change process is genuine rather than a statistical artifact.

As with most prior clinical studies of this population, sample attrition posed an important threat to the validity of conclusions. The low rate of partner contact (61% at follow-up) is similar to rates obtained in previous work (Dunford, 2000; Edleson & Syers, 1990; Saunders, 1996), even with extensive efforts to minimize attrition (Gondolf, 1997). It remains possible that clients whose partners did not provide complete follow-up data benefited less than others from the treatments. It is unlikely, however, that attrition could fully account for the observed reductions in relationship aggression. Participants whose partners were lost to follow-up did not have higher rates of pre-treatment aggression, suggesting that they were not more problematic cases at the outset. In addition, the finding of abuse reduction over time with no differences between treatments was corroborated by criminal justice outcomes and client self-report data.

The follow-up period for partner reports was relatively brief (6 months after treatment and approximately 12 months after intake). However, a recent multisite study of abuser treatments found that almost half of all men who reassaulted their partners did so within the first 3 months of treatment. By 9 months after intake, first-time reassaults decreased to 3% to 4% per quarterly assessment (Gondolf, 1997). Therefore, the small percentage of new abuse incidents identified through a longer follow-up would not likely have altered the study conclusions.

Several factors might explain the lack of significant differences in outcome for the two treatment conditions. There were a small number of high-rate reoffenders (approximately 5% of the sample), a sizeable proportion of cases with no repeat violence, and a "middle" group who engaged in one or two physically aggressive behaviors at posttreatment or follow-up. It is likely that factors such as arrest, successful prosecution, and probation monitoring (Murphy et al., 1998) could result in a no-violence outcome for many cases with little or no treatment. For example, a recent large-scale investigation found a violence recidivism rate among no-treatment controls of about 30% during the year after treatment (Dunford, 2000). High-rate reoffenders, on the other hand, may be hard to reach with brief group interventions because of the severity of their problem behavior, a lack of motivation to change, or serious personality problems that impede treatment (Dutton, Bonarchuk, Kropp, Hart, & Ogloff, 1997; Hamberger & Hastings, 1990; Weisz, Tolman, & Saunders, 2000). Changes in life situation, relationship status, and other partner-violence risk factors further complicate efforts to detect treatment effects. About half of the participants were separated from the abuse victim at program intake. Some reunite during treatment, others split up, and others fluctuate in their contact and relationship status over time. Although risk of repeat violence is present in all cases, opportunity varies considerably with the



degree and nature of ongoing relationship contact (Fals-Stewart, Lucente, & Birchler, 2002). Changes in comorbid problems such as depression or substance abuse may exert further uncontrolled effects on outcome.

In addition to client heterogeneity, common factors across treatment conditions may be responsible for the similarity in treatment efficacy. Both treatments employed strategies to promote group cohesion and therapeutic alliance, were conducted by the same pool of therapists, and used motivational techniques to enhance treatment attendance (Taft, Murphy, Elliott, & Morrel, 2001). Both groups also received basic instruction in taking time-outs during relationship conflicts, a potentially helpful behavioral intervention. As noted in the introduction, several prior studies of domestic violence interventions have likewise found little or no difference in efficacy for diverse group treatment approaches (Brannen & Rubin, 1996; Edleson & Syers, 1990; O'Leary, Heyman, & Neidig, 1999; Saunders, 1996).

Clinical impressions and treatment adherence data suggest two explanations for the unexpected findings regarding the lack of added benefit from structured CBT interventions. The first is that a reliance on didactic interventions in CBT may have inadvertently reduced opportunities for constructive group behaviors, such that men in the ST groups had more opportunity to engage in supportive and empathic conversation with each other. This informal practice may have shaped communication skills and enhanced self-efficacy for handling relationship difficulties even in the absence of direct communication training. A second possible explanation involves incomplete or inadequate implementation of CBT principles. Despite direct instruction and role-play activities, CBT participants had considerable difficulty implementing anger-control strategies and communication skills such as active listening and emotional communication, even with frequent coaching from the therapists. The fact that only 4 of the 16 CBT sessions were allocated to communication skills training may have limited participants' ability to become proficient in abuse-alternative behaviors. In addition, the adherence analyses revealed that therapists engaged in CBT-related therapist activities during only about one-third of session intervals, which raises the possibility that the level of skills training and practice may have been inadequate to produce generalized change in target behaviors. It is interesting to note that one of the earliest intervention trials for abusive men focused almost exclusively on relationship enhancement via communication training, showing reduced criminal recidivism relative to untreated controls (Waldo, 1988). Future studies of CBT for partner-violent men need to examine factors such as homework compliance, change in target cognitions, and mastery of behavioral skills. It remains unclear whether changes in these intermediate (proximal) targets does not lead to reductions in abusive behavior, or whether the current CBT intervention did not sufficiently alter these intermediate treatment targets.

In conclusion, the findings indicate substantial pre- to posttreatment reductions in abusive behavior that were maintained at a subsequent follow-up assessment for participants in both CBT and ST. Although the treatments

were highly discriminable, no differences were observed on the major outcome measures of abusive behavior. More detailed analyses of CBT interventions and further development of CBT approaches for this population are sorely needed. Recent efforts to use motivational enhancement techniques to address client resistance to change (Kistenmacher & Weiss, 2001; Musser, Murphy, & Taft, 2001) and dialectical behavior therapy to address affective dysregulation and complex problems in partner-violent men (Fruzzetti & Levensky, 2000) provide promising alternatives to the standard group interventions in the field.

## References

- Babcock, J. C., & LaTaillade, J. (2000). Evaluating interventions for men who batter. In J. Vincent & E. Jouriles (Eds.), *Domestic violence: Guidelines for research-informed practice* (pp. 37–77). Philadelphia: Jessica Kingsley.
- Babor, T. F., de la Fuente, J. R., Saunders, J., & Grant, M. (1992). *AUDIT: The Alcohol Use Disorders Identification Test: Guidelines for use in primary health care*. Geneva, Switzerland: World Health Organization.
- Bandura, A. (1986). *Social foundations of thought and action: A social and cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Begun, A. L., Murphy, C. M., Bolt, D., Weinstein, B., Strodthoff, T., Short, L., & Shelley, G. (in press). Characteristics of the Safe at Home instrument for assessing readiness to change intimate partner violence. *Research on Social Work Practice*.
- Brannen, S. J., & Rubin, A. (1996). Comparing the effectiveness of gender-specific and couples groups in a court-mandated spouse abuse treatment program. *Research on Social Work Practice, 6*, 405–424.
- Demo, D. H. (1985). The measurement of self-esteem: Refining our methods. *Journal of Personality and Social Psychology, 48*, 1490–1502.
- Dunford, F. W. (2000). The San Diego Navy experiment: An assessment of interventions for men who assault their wives. *Journal of Consulting and Clinical Psychology, 68*, 468–476.
- Dutton, D. G. (1986). The outcome of court-mandated treatment for wife assault: A quasi-experimental evaluation. *Violence and Victims, 1*, 163–175.
- Dutton, D. G., Bonarchuk, M., Kropp, R., Hart, S. D., & Ogloff, J. P. (1997). Client personality disorders affecting wife assault post-treatment recidivism. *Violence and Victims, 12*, 37–50.
- Edleson, J. L., & Syers, M. (1990). Relative effectiveness of group treatments for men who batter. *Social Work Research and Abstracts, 26*, 10–17.
- Fals-Stewart, W., Lucente, S. W., & Birchler, G. R. (2002). The relationship between the amount of face-to-face contact and partners' reports of domestic violence frequency. *Assessment, 9*, 123–130.
- Faulkner, K., Stoltenberg, C. D., Cogen, R., Nolder, M., & Shooter, E. (1992). Cognitive-behavioral group treatment for spouse abusers. *Journal of Family Violence, 7*, 37–55.
- Fleming, J. S., & Courtney, B. E. (1984). The dimensionality of self-esteem: II. Hierarchical facet model for revised measurement scales. *Journal of Personality and Social Psychology, 46*, 404–421.
- Fruzzetti, A. E., & Levensky, E. R. (2000). Dialectical behavior therapy for domestic violence: Rationale and procedures. *Cognitive and Behavioral Practice, 7*, 435–447.
- Gondolf, E. W. (1997). Patterns of reassault in batterer programs. *Violence and Victims, 12*, 373–387.
- Gondolf, E. W., & White, R. J. (2000). "Consumers" recommendations for batterers programs. *Violence Against Women, 6*, 198–217.

- Hamberger, L. K., & Hastings, J. E. (1988). Skills training for treatment of spouse abusers: An outcome study. *Journal of Family Violence*, 3, 121–130.
- Hamberger, L. K., & Hastings, J. E. (1990). Recidivism following spouse abuse abatement counseling: Treatment program implications. *Violence and Victims*, 5, 157–170.
- Haynes, S. N., Chavez, R. E., & Samuel, V. (1984). Assessment of marital communication and distress. *Behavioral Assessment*, 6, 315–321.
- Holtzworth-Munroe, A., & Hutchinson, G. (1993). Attributing negative intent to wife behavior: The attributions of maritally violent versus nonviolent men. *Journal of Abnormal Psychology*, 102, 206–211.
- Jennings, J. L. (1987). History and issues in the treatment of battering men: A case for unstructured group therapy. *Journal of Family Violence*, 2, 193–213.
- Jennings, J. L., & Murphy, C. M. (2000). Male-male dimensions of male-female battering: A new look at domestic violence. *Psychology of Men and Masculinity*, 1, 1–9.
- Kistenmacher, B., & Weiss, R. L. (2001, July). *Motivational interviewing as a mechanism for change in men who batter*. Presented at the 7th International Family Violence Research Conference, Portsmouth, NH.
- Margolin, G., John, R. S., & Gleberman, L. (1988). Affective responses to conflictual discussions in violent and nonviolent couples. *Journal of Consulting and Clinical Psychology*, 56, 24–33.
- Miller, W. R., & Rollnick, S. (1991). *Motivational interviewing: Preparing people to change addictive behavior*. New York: The Guilford Press.
- Morrel, T. M., & Murphy, C. M. (1996, November). *The role of self-esteem and self-efficacy in cognitive-behavioral treatment for domestic abusers*. Paper presented at the annual convention of the Association for Advancement of Behavior Therapy, New York.
- Murphy, C. M., Musser, P. H., & Maton, K. I. (1998). Coordinated community intervention for domestic abusers: Intervention system involvement and criminal recidivism. *Journal of Family Violence*, 13, 263–284.
- Murphy, C. M., & Scott, E. (1995). *Cognitive behavior therapy for domestically assaultive individuals: A treatment manual*. Unpublished manuscript, University of Maryland, Baltimore County.
- Musser, P., Murphy, C., & Taft, C. (2001, July). *Motivational interviewing and domestic abuse perpetrators*. Presented at the 7th International Family Violence Research Conference, Portsmouth, NH.
- Novaco, R. W. (1975). *Anger control*. Lexington, MA: DC Heath.
- O'Leary, K. D., Heyman, R. E., & Neidig, P. N. (1999). Treatment of wife abuse: A comparison of gender-specific and conjoint approaches. *Behavior Therapy*, 30, 475–505.
- Palmer, S. E., Brown, R. A., & Barrera, M. E. (1992). Group treatment program for abusive husbands: Long-term evaluation. *American Journal of Orthopsychiatry*, 62, 276–283.
- Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist*, 47, 1102–1114.
- Rosenbaum, A., & O'Leary, K. D. (1986). The treatment of marital violence. In N. S. Jacobson & A. S. Gurman (Eds.), *Clinical handbook of marital therapy* (pp. 385–405). New York: The Guilford Press.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Saunders, D. G. (1996). Feminist-cognitive-behavioral and process-psychodynamic treatments for men who batter: Interactions of abuser traits and treatment models. *Violence and Victims*, 11, 393–414.
- Saunders, D. G., & Hanusa, D. (1986). Cognitive-behavioral treatment of men who batter: The short-term effects of group therapy. *Journal of Family Violence*, 1, 357–372.
- Sonkin, D. J., & Durphy, M. (1989). *Learning to live without violence*. Volcano, CA: Volcano Press.

- Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The revised conflict tactics scales (CTS2). *Journal of Family Issues, 17*, 283–316.
- Taft, C. T., Murphy, C. M., Elliott, J. D., & Morrel, T. M. (2001). Attendance enhancing procedures in group counseling for domestic abusers. *Journal of Counseling Psychology, 48*, 51–60.
- Vivian, D., & Langhinrichsen-Rohling, J. (1994). Are bi-directionally violent couples mutually victimized? A gender-sensitive comparison. *Violence and Victims, 9*, 107–124.
- Waldo, M. (1988). Relationship enhancement counseling groups for wife abusers. *Journal of Mental Health Counseling, 10*, 37–45.
- Weisz, A. N., Tolman, R. M., & Saunders, D. G. (2000). Assessing the risk of severe domestic violence: The importance of survivors' predictions. *Journal of Interpersonal Violence, 15*, 75–90.
- Yalom, I. D. (1995). *The theory and practice of group psychotherapy* (4th ed.). New York: Basic Books.

RECEIVED: January 23, 2002

ACCEPTED: July 30, 2002

## Treatment of Wife Abuse: A Comparison of Gender-Specific and Conjoint Approaches

K. DANIEL O'LEARY

RICHARD E. HEYMAN

PETER H. NEIDIG

*State University of New York at Stony Brook*

Seventy-five intact, volunteer couples were assigned to either a gender-specific or a conjoint 14-week group treatment for psychological and physical aggression. Participants from both treatments significantly reduced their psychological and physical aggression, at both posttreatment and 1-year follow-up. During treatment, husbands reduced their psychological aggression by 47%, their moderate physical aggression by 55%, and their severe physical aggression by 51%. Although two-thirds of the husbands maintained cessation of severe aggression during the year following treatment, only one-fourth of the husbands were violence-free. Very similar cessation and maintenance rates were obtained for wives. Significant improvements at posttreatment and follow-up were also found for both spouses' marital adjustment, husbands' taking responsibility for aggression, and wives' depression. No differential effect of treatment type was found, except that, as predicted, husbands in conjoint treatment improved more on marital adjustment. Neither form of treatment was superior to the other in terms of safety and effectiveness for volunteer, intact, and physically aggressive couples.

Research on wife abuse has moved from almost nonexistence to national prominence, with research agendas on family violence being offered by the National Research Council (Crowell & Burgess, 1996) and the American Psychological Association (APA, 1996). Despite the explosion of research on the prevalence and correlates of partner abuse (Straus & Gelles, 1990), treatment outcome research is less well developed. Research has only recently begun to compare efficacy across theoretical approaches, such as between cognitive behavioral and psychodynamic treatment (e.g., Saunders, 1996).

Most treatment outcome research is conducted with court-mandated samples and often with men who are not in intact relationships (Hamberger, 1996; Rosenbaum, Gearan, & Ondovic, 1997). However, a far larger popula-

Address correspondence to K. Daniel O'Leary, Ph.D., Department of Psychology, State University of New York at Stony Brook, NY 11794-2500.

tion of aggressive men exists: 50% to 65% of clinically martially discordant men are physically aggressive (Holtzworth-Munroe et al., 1992; O'Leary, Vivian, & Malone, 1992), with approximately half of these couples reporting severe physical aggression (e.g., hitting or beating). However, treatment recommendations formulated for court-mandated men are typically applied to all men—even those who have committed a single act of aggression (e.g., McMahon & Pence, 1996). Given that the majority of court-mandated men are single, separated, or divorced (Rosenbaum et al.), clearly more must be known about treating aggressive, married men before making policy decisions regarding their treatment. This is especially important, given that 90% of those in aggressive relationships who seek generic marital therapy do not report aggression as a presenting problem (Ehrensaft & Vivian, 1996; O'Leary et al., 1992). The aim of this study was to provide a comparison of the effectiveness of two treatment approaches, gender-specific treatment (GST; i.e., men's and women's groups) and conjoint treatment focusing on the reduction of psychological and physical aggression, in a self-referring, maritally intact, physically aggressive sample.

We do not advocate conjoint treatment for all couples in physically abusive relationships, and we sought to test its utility under the following conditions: (a) the couple is intact; (b) the violence is not severe enough to elicit serious injury or substantial fear; and (c) the participants either acknowledge aggression as a problem or are willing to participate in a program that makes aggression the primary target of treatment. Nevertheless, conjoint treatment is controversial for several reasons. First, some professionals believe that physically victimized wives may not feel comfortable expressing themselves in conjoint sessions for fear of reprisal (Ganley, 1981). Second, some maintain that conjoint treatment could lead wives to believe that they were partly responsible for their husband's aggression (see O'Leary, 1996; McMahon & Pence, 1996, for a debate on this issue). Third, some therapists fear that wives' in-session comments could be a stimulus for their husbands' later retribution and violence. These are empirical questions addressed in the current study.

### *GST: Background and Rationale*

GST is the dominant form of treatment for wife abuse, which typically melds a feminist theoretical framework with a cognitive-behavioral approach. Feminists (cf. Yllö, 1993) posit that aggression is completely under men's control; men use psychological and physical aggression for power and control; and men should be the sole targets for intervention. Feminists believe that male aggression has a gender-based function (i.e., to preserve male power and control) and highlight the greater psychological and physical impact of physical violence on women (e.g., Cascardi, Langhinrichsen, & Vivian, 1992; Stets & Straus, 1990). Treatment implications are thus that (a) men must take sole responsibility for past abuse and for stopping abuse; (b) men must recognize male power and control at the societal and family level and acknowledge their abusive beliefs about male power that support

aggression in the home; and (c) men must be treated in separate groups, because limiting treatment to men emphasizes their sole responsibility for abuse cessation. Cognitive-behavioral interventions emphasize "learned patterns of self-produced self-talk, overt behavior, and physiological arousal of the individual batterer, even in the absence of external behaviors. As such, the individual batterer is clearly responsible for his behavior" (Hamberger & Lohr, 1989, p. 69).

Most GST programs for men accept that there are a broad array of variables associated with partner abuse (e.g., men's control of women, poor impulse control, and defective communication), and thus adopt an eclectic approach (Rosenbaum & Maiuro, 1989). Some of the guiding principles of such programs are "Abusers do not enjoy being abusive [the typical batterer is unhappy with his life and his behavior]. Batterers often have negative attitudes toward women, which contribute to the occurrence of aggression in marital relationships. Abusers generally lack nonviolent alternatives for expressing themselves or for achieving desired goals within their marital relationships. Marital aggression is an individual's problem, most commonly the man's. Marital discord is a couple's problem" (Rosenbaum & Maiuro, pp. 166-176).

GST models are compelling for the treatment of men who beat and intimidate their partners, and we recommend such treatments for those cases (Rosenbaum & O'Leary, 1986; Vivian & Heyman, 1996). On the other hand, their applicability to volunteer treatment seekers in intact relationships with moderate levels of physical aggression is not clear.

### *Conjoint Treatment: Background and Rationale*

Early approaches using conjoint treatment (Geffner, Mantooth, Franks, & Rao, 1989; Lindquist, Telch, & Taylor, 1983; Neidig & Friedman, 1984) followed an eclectic model based on cognitive, social learning, and family systems theories. Such programs, although not completely antithetical to feminist approaches, tended to emphasize the bi-directional, interactional factors in partner/marital violence and to de-emphasize labeling men as perpetrators and women as victims.

Several findings about the marital context of physical aggression help put the rationale for conjoint treatment in perspective. First, aggression in intimate relationships typically occurs in the context of an argument between partners (Cascardi & Vivian, 1995; Stamp & Sabourin, 1995). Second, although the impact of husband-to-wife aggression is perceived by the victim as negative (Cascardi et al., 1992), few women presenting for couples therapy are fearful of remaining with their husbands or of participating in a treatment with them. Third, in marital clinic samples, aggression is often mutual (Vivian & Langhinrichsen-Rohling, 1994) and is in self-defense in less than 20% of the cases (Cascardi & Vivian). While recognizing the differential impact of aggression on women, conjoint treatment proponents believe men's violence is most likely to cease when both partners adopt a nonviolent philosophy (e.g., Vivian & Heyman, 1996). Fourth, Rathus, O'Leary, and Meyer (in

press), studying men, found no differences in dominance/isolation behaviors between physically aggressive men and nonaggressive men in discordant, intact marriages. Tolman (1999), studying women, reported similar findings in an aggressive, non-court-mandated sample. In short, power and control tactics (the predominant focus of most GST programs) may characterize some, but clearly not all men, and these tactics are not confined to aggressive marriages. Finally, because marital discord is one of the strongest risk factors for partner abuse (Pan, Neidig, & O'Leary, 1994a), advocates of conjoint approaches believe that reduction of marital discord should make partner abuse less likely.

### *Prior Comparison of GST and Conjoint Treatment*

Brannen and Rubin (1996) compared GST and conjoint approaches with court-mandated men ( $N = 49$ ) referred by the criminal justice system in San Antonio, TX. Two-thirds of the sample was Hispanic; the mean family income was \$19,750. The conjoint treatment program used in the current study and in Brannen and Rubin's treatment evaluation were both variants of that described in Neidig and Freidman (1984). Both treatments were associated with significant reductions in physical aggression (as assessed by wife reports at posttreatment and at 6-month follow-up). Brannen and Rubin reported impressive violence cessation rates (92%, as reported by the wives), given that success rates typically range from 53% to 85% (Edelson & Tolman, 1992). For couples with a history of alcohol problems, the conjoint approach was superior. Furthermore, of the 7 dropouts prior to treatment, 6 were from the GST. The successful outcome may have been influenced by several factors, such as the associated surveillance from the probation department and a follow-up rate of 62%. In brief, the Brannen and Rubin study provides some indication that a conjoint intervention can be used successfully (albeit in conjunction with probation and associated surveillance).

As noted earlier, the majority of physically aggressive couples in the United States are not mandated to treatment. Indeed, because 44% of all young married (O'Leary et al., 1989) and 50% to 67% of maritally distressed (O'Leary et al., 1992) couples have relationships characterized by physical aggression, it would be practically impossible to mandate all such individuals to treatment. The majority of couples in physically aggressive relationships who get services do so through mental health clinics, marital clinics, and private practitioners.

### *Study Design*

Treatment studies of wife abuse present difficulties for straightforward treatment versus untreated control group designs (Fagan, 1996). Although untreated control groups might best allow one to know about the natural course of a problem and/or use of community services, there are ethical problems in withholding treatment for long periods because of potential repeated aggression and injury (Caesar & Hamberger, 1989; Hamberger & Hastings, 1988), and they have rarely, if ever, been used in published violence research



(Edelson & Tolman, 1992). Hence, in this initial evaluation of our treatments, we decided that the most prudent course of action would be to test conjoint treatment against the more widely used treatment, GST. This comparison seemed reasonable because we had differential predictions about the treatment outcome.

Because of the widespread prevalence of physical aggression in distressed couples who are *not* court-mandated to treatment, and because of the dearth of studies on whether GSTs or conjoint treatments were more effective with that subpopulation, we recruited volunteer, maritally intact couples (a) who reported two or more acts of husband-to-wife physical aggression in the past year; but (b) whose physical aggression did not produce injuries necessitating medical attention. An additional inclusion criterion was that wives, when interviewed separately, had to report that they would feel comfortable being in a conjoint treatment with their husbands.

We hypothesized that, although both treatments would decrease partner aggression, spouses in conjoint treatment would improve their marital adjustment and communication problems significantly more than those in GST. We predicted that increases in men's marital satisfaction would be associated with decreases in their physical aggression. Given the expected differences in posttreatment marital processes between the groups, we expected less husband-to-wife aggression in the conjoint group at follow-up than in GST.

We tested prevalent beliefs about relative risks and benefits of conjoint and gender-specific approaches. As noted above, some GST proponents believe that any program that treats husbands and wives conjointly is inherently more risky than programs that separate spouses. However, we expected that both approaches would be safe and effective in halting abuse. Thus, we tested the following additional predictions that emanate from critiques and concerns about conjoint treatment: (a) women will be more fearful of expressing themselves in conjoint treatment than GST; (b) conjoint treatment sessions will result in more post-session husband-to-wife violence than GST sessions; (c) GST will result in less male dominance/isolation behaviors at posttreatment and follow-up than conjoint treatment; (d) wives in GST will report less fear and depressive symptomatology at posttreatment and follow-up than those in conjoint treatment; and (e) husbands and wives in GST, compared to those in conjoint treatment, will be more likely to attribute responsibility for husband-to-wife violence to the husband and not the wife.

## Method

### *Participants*

Participants responded to newspaper announcements offering free therapy for spouses whose "arguments led to throwing, pushing, shoving, etc." Interviewers telephone-screened each partner privately and confidentially with the Conflict Tactics Scale (Straus, 1979). As the primary focus was on husbands' aggression, we accepted only couples with repeated acts of husband-to-

wife physical aggression. Our inclusion criteria of two acts of aggression was because single acts of aggression are the least likely to persist across time (Malone & O'Leary, 1995); and we thought that the program's explicit focus on aggression required more than a single, isolated instance of aggression by men, especially since approximately 40% of newly married men *and* women report such physical aggression against their partners (O'Leary et al., 1989). Thus, to qualify for an in-person assessment, either partner had to report two or more acts of husband-to-wife aggression during the past year. Furthermore, to ensure group cohesion in the men's groups, we required that all men report at least one act of husband-to-wife aggression during the telephone screen.

Approximately 800 phone inquiries were received. Of those who were screened out over the telephone,<sup>1</sup> the most frequent reason was absence of physical aggression; the second most frequent reason was that the partners were not married. In addition, in approximately 10% of the calls, women reported that they were the sole or primary aggressor. As detailed in Brown, O'Leary, and Feldbau (1997), approximately 40% of the total callers met the telephone screen criteria for inclusion. About two-thirds of these callers declined to participate further—approximately one third said that they were not interested (with no further explanation), one third had scheduling problems, and approximately 20% wanted individualized, not group, treatment (for which they were given referrals).

Ninety-four couples completed the in-person intake assessment. (Five more couples had only one spouse complete the assessment, and were thus eliminated.) The following criteria determined inclusion in the treatment study: (a) both spouses were willing to be randomly assigned to either mode of treatment; (b) the wife did not report sustaining injuries that required medical attention; (c) the wife, during a private interview, reported feeling comfortable being assigned to conjoint treatment; (d) the wife was not afraid of living with her husband; (e) the husband did not meet criteria for alcohol dependence on the Structured Clinical Interview for *DSM-III-R* (SCID); (f) neither spouse reported psychotic symptoms nor met criteria for a diagnosis of current psychopathology severe enough to interfere with successful participation in the group (e.g., bipolar disorder).

We had anticipated the need to screen out many couples at the initial in-person interview due to severe injuries or fear of the partner. However, we did not have large-scale exclusions from treatment. Five couples were disqualified: one woman had endured severe injury; one husband was bipolar and his mania was not psychiatrically stabilized; one couple came drunk to the assessment, and both partners met criteria for current alcohol dependence; one wife was suicidal; and it became clear that one wife was the primary aggressor. Fourteen couples withdrew from the study while waiting for group

<sup>1</sup> Couples who were screened out were provided with referrals to local service providers, including battered women shelters and county batterers programs.

assignment (including two couples who separated while waiting for the next group). Thus, 75 couples were assigned to treatments. Of those, 5 (7%) never attended a group session, 33 (44%) came to at least one session but no more than nine sessions (and were considered dropouts), and 37 (49%) completed treatment. Finally, 31 of the 37 completers (84%) participated in the follow-up. No couples who completed the treatment programs separated during the treatment, but two couples separated in the year following treatment. All data except marital adjustment were collected for these couples.

Spouses assigned to the conjoint and GST groups were equivalent on the following variables, as tested with a series of MANOVAs: pretreatment levels of psychological, mild physical, and severe physical aggression; fear of partner; depressive symptomatology; marital adjustment; education; and family income. Husbands' age was significantly different between the forms of treatments  $F(1, 65) = 5.78, p < .05$ , with those assigned to conjoint treatment ( $M = 39.33, SD = 10.18$ ) significantly older than those assigned to GST ( $M = 35.27, SD = 6.67$ ).

Brown et al. (1997) made a comparison of those who completed treatment with those who dropped out. The most frequently cited reasons for dropout by participants were treatment-related issues (e.g., dropouts reported that the group setting did not enable them to address their individual couple issues). The mean number of sessions attended by those who dropped out of the program was 4.2 ( $SD = 2.66$ ); for those who completed the program,  $M = 12.31$  ( $SD = 1.29$ ).

For this article, we conducted three logistic regression analyses to predict differential dropout between conjoint treatment and GST. The dependent variable in each analysis was dropout/completer status. The independent variables were interactions between treatments (conjoint or GST) and the following: (analysis 1) husband's psychological, mild physical and severe physical aggression; dominance/isolation; wives' fear; wives' depression; (analysis 2) husbands' and wives' marital adjustment; and (analysis 3) husbands' and wives' education and age; family income. None of the three logistic regression models significantly predicted dropout. All treatment type  $\times$  variable interactions were non-significant, except husbands' and wives' ages, with younger couples being more likely to drop out of GST: (Husbands' Age: Wald's coefficient [ $Z$ ] = 4.48,  $p < .05$ ; Wives' Age:  $Z = 3.86, p < .05$ ).

### *Demographics of Treatment Completers*

We served a suburban sample on Long Island (Suffolk County), NY. Of treatment completers, one wife was African American and one wife was Hispanic; one husband was African American; the remainder of the completers were Caucasian. The mean annual family income was \$51,454 ( $SD = \$23,380$ ), similar to the Suffolk County average (i.e., median family income was \$53,247 in the 1990 census). The average age of the women was 36.24 ( $SD = 7.86$ ), and the average age of the men was 38.40 ( $SD = 8.79$ ). The average educational level of women was 13.54 ( $SD = 1.88$ ) and men was 13.56 years ( $SD = 1.80$ ).

### Measures

The seven self-report measures described below were administered at pretreatment, posttreatment, and follow-up.

1. *Modified Conflict Tactics Scale* (MCTS; Pan, Neidig, & O'Leary, 1994b). The MCTS adds six additional items to Straus' (1979) original 18-item Conflict Tactics Scale (CTS). The 24-item self-report inventory assesses the frequency of a variety of functional (e.g., calmly discussing a problem), verbally abusive (e.g., insults or swearing), and physically abusive (e.g., hitting) conflict tactics. The CTS has been used in national surveys of the prevalence of marital aggression (Straus & Gelles, 1990; Straus, Gelles, & Steinmetz, 1980), and it is the most widely used measure of physical abuse in intimate relationships.

Our psychological abuse scale, guided by the Pan et al. (1994b) factor analysis, comprised the following items: insult or swear; sulked and/or refused to talk about it; stomped out of the room, house, or yard; did or said something to spite spouse; threatened to leave the marriage; threatened to do things like withhold money, have an affair, etc.; threaten to hit or throw something at spouse; threw, smashed, hit or kicked something; drove recklessly to frighten spouse. The scale was internally consistent in this sample (pretreatment Cronbach's  $\alpha$ s for self- and partner-reports ranged from .75 to .87).

We used the standard (Straus, 1979), theoretically based, a priori definitions for mild aggression (controlled spouse physically; threw something at partner; pushed, grabbed, or shoved; slapped) and severe aggression (kicked, bit, hit with a fist; choked; beat up; physically forced sex; threatened with gun or knife; used gun or knife). Each item was rated on a frequency scale for its occurrence during the past year: never, 1 time, 2 times, 3 to 5 times, 6 to 10 times, 11 to 20 times, more than 20 times. Following Straus, each response that provided ranges (e.g., 3 to 5 times) were coded at the midpoint, generating a scale of 0, 1, 2, 4, 8, 16, and 25. Spouses were asked about their own and their spouses behavior. Thus, each spouse received four physical aggression scores for each administration of the MCTS, each of which was a sum of the items that constituted that scale: husband-to-wife (mild), husband-to-wife (severe), wife-to-husband (mild), and wife-to-husband (severe).

The MCTS was administered at pretreatment and follow-up with standard instructions to report on the past year. The scale was also administered at the first and last treatment session with instructions to report on the prior 14 weeks (i.e., a time-frame equivalent to the length of treatment).

The levels of husbands' aggression, using the 14-week and 1-year pretreatment reports, were as follows: pushed, grabbed, or shoved (79% [14 week], 100% [1 year]); tried to control physically (70%, 91%); threw something at wife (52%, 75%); slapped (36%, 64%); kicked, bit, or hit with fist (36%, 56%); choked or strangled (21%, 36%); beat up (24%, 25%); physically forced sex (3%, 11%); threatened with knife or gun (15%, 17%); used knife or gun against wife (0%, 0%). Overall, there were 19.31 mild and 3.34 severe acts of husband-to-wife physical aggression reported in the past year.

Interspousal agreement for husband-to-wife-psychological aggression was not significant at either 1 year pretreatment or at 14 week pretreatment. On the other hand, agreement for physical aggression was significant for three of the four instances assessed. For 1 year pretreatment ( $n = 36$ ), agreement regarding physical aggression was as follows: mild,  $r = .20$ , ns; severe,  $r = .30$ ,  $p < .05$ . For 14 week pretreatment ( $n = 37$ ), agreement was as follows: mild,  $r = .37$ ,  $p < .05$ ; severe,  $r = .29$ ,  $p < .05$ .

Interspousal agreement for wife-to-husband aggression was significant for both 1 year pretreatment and 14 weeks pretreatment. For 1 year pretreatment ( $n = 36$ ), agreement was as follows: psychological,  $r = .45$ ,  $p < .001$ ; mild,  $r = .57$ ,  $p < .001$ ; severe,  $r = .49$ ,  $p < .001$ . For 14 week pretreatment ( $n = 37$ ), agreement was as follows: psychological,  $r = .52$ ,  $p < .001$ ; mild,  $r = .52$ ,  $p < .001$ ; severe,  $r = .60$ ,  $p < .001$ . The agreement levels for wife-to-husband aggression were in the moderate level range (.45 to .60), and were similar to the reliabilities generally obtained in the partner abuse literature with clinical samples using correlations of aggression scores (Cantos, Neidig, & O'Leary, 1994; Jouriles & O'Leary, 1985). However, the agreement for reports of husband-to-wife aggression were lower than those obtained by us and by others with community samples (e.g., Moffitt et al., 1997). One reason for the lower values obtained here is that we divided the physical aggression into mild and severe aggression, and husband-to-wife agreement has not been reported in other studies with this distinction. In addition, because of the often discussed underreporting of partner abuse by men (Heyman & Schlee, 1997), several options are often suggested, namely, use of the wife's report or use of the higher score by either partner (either/or method). We chose the second method, although we conducted our analyses both ways, and the results did not differ.

2. *Dominance/Isolation Scale* (Tolman, 1989). Fourteen items that had the highest factor loadings (i.e., above .60) on the dominance and isolation factor of the Psychological Maltreatment of Women Scale comprised the Dominance/Isolation Scale in this study. Among the items were the following: "ordered me around"; "monitored my time and made me account for where I was"; "restricted my use of the car." Respondents indicated the frequency with which their partners engaged in that behavior during the past year. In this sample, Cronbach's  $\alpha$  was .84 (husbands' reports about wives) and .92 (wives' reports about husbands). Frequency ratings ranged from 0 (never) to 4 (very frequently); scores could range from 0 to 56.

3. *Fear of Spouse*. We employed a face valid, 11-item fear scale previously used by O'Leary and Curley (1986). Question one asked women to rate, from 0 (not at all) to 6 (highly fearful), their fear of their husbands. The remaining items asked whether the women were fearful (on a true/false basis) that their husbands would do any of 10 threatening or abusive behaviors (from getting upset to physically injuring me). The range for this measure is 0 to 16. Cronbach's  $\alpha$  was .72 (men) and .72 (women) for the fear scale administered in the pretreatment packet.

4. *Attribution of Responsibility Items.* Four items on causal responsibility for physical aggression were used to assess attribution for responsibility. Two stems ("When I am physically aggressive" and "When my partner is physically aggressive") are followed by two response questions each ("It is caused by things my partner says or does," "It is caused by things I say or do.") Ratings for each range from 1 (strongly agree) to 5 (strongly disagree). A skip out response was provided if respondent reported no aggression occurred during the specified time-frame.

5. *Beck Depression Inventory* (BDI; Beck, Steer, & Garbin, 1988). BDI is a measure of depressive symptomatology, consisting of 21 items, each corresponding to a specific category of symptoms and attitudes. The inventory has been shown to be reliable and valid. The BDI consistently correlates with clinical ratings of depression, and has been used in scores of treatment outcome studies. Cronbach's  $\alpha$  was .84 (men) and .83 (women) for the BDI administered in the pretreatment packet.

6. *Dyadic Adjustment Scale* (DAS; Spanier, 1976). The DAS is a 32-item self-report inventory designed to measure the severity of relationship discord in intimate dyads. Scores range from 0–151, with higher values indicating more favorable adjustment; a score below 98 is traditionally considered to indicate clinical marital distress (e.g., Jacobson & Truax, 1991). The DAS has high convergent validity with other measures of marital adjustment and satisfaction (e.g., Heyman, Sayers, & Bellack, 1994). More than 1,000 published studies have used the DAS. In this sample, Cronbach's  $\alpha$  was .89 (men) and .90 (women) for the DAS administered in the pretreatment packet. As noted below, some follow-ups were conducted by telephone. Because the DAS is cumbersome to administer over the telephone due to its length and many different rating scales used, we used the validated, telephone version (Krokoff, 1989) of the Marital Adjustment Test (MAT; Locke & Wallace, 1959). MAT scores were then converted to DAS scores with a formula published by Crane, Allgood, Larson, and Griffin (1990).

7. *Spouse Verbal Problems Checklist* (SVPC; Haynes, Chavez, & Samuel, 1984). The original SVPC is a 27-item self-report inventory exploring the communication deficits of the respondent's partner. Haynes et al. found evidence of convergent validity, with the SVPC being strongly associated with observers' coding of marital conflict. Vivian and Malone (1997) conducted a factor analysis of the SVPC on 327 couples who sought marital therapy between 1980 and 1990, and found that 18 items loaded on a factor. We administered this 18-item version to participants in the present study. Cronbach's  $\alpha$  was .87 (men) and .88 (women) for the SVPC administered in the pretreatment packet.

*SCID, nonpatient edition* (Spitzer, Williams, Gibbons, & First, 1992). The SCID, with the supplementary PTSD module, was administered to each spouse privately during the pretreatment interview only. SCID interviews were conducted by graduate students in clinical psychology and a bachelor's level research assistant. SCID training was conducted until a minimum of

95% interrater agreement was achieved. Twenty percent of the SCID audiotapes were randomly assigned to a second rater for reliability scoring. Potential participants were screened out if the husband met criteria for current alcohol dependence (interrater agreement = 100%) or met any of the nine psychotic screen items (interrater agreement = 100%).

*Fear and/or aggression due to treatment sessions.* Before each treatment session, participants completed a brief questionnaire, which included items on fear during the last session ("Have you been afraid to express your thoughts or feelings during the last treatment session?" on a 1 [not at all afraid] to 5 [afraid all the time] scale) and treatment precipitating violent arguments ("Has anything discussed in last week's treatment session led to a serious physical argument?"). To account for missed sessions, we counted the percent of sessions that each wife attended in which she indicated that aggression had occurred.

*Consumer satisfaction.* Five items about satisfaction with the treatment received were included in the posttreatment and follow-up packets: "How do you feel about: (a) Your interest and involvement in the program; (b) How relevant the program was for you; (c) The personality of the therapist(s); (d) Therapists's skill and competence; and (e) How much your own concerns/goals were met by the program." Ratings were on a 1 (extremely negative) to 9 (extremely positive) scale.

### *Procedure*

Participants who met telephone screening criteria were scheduled for a 3 hour, in-person pretreatment assessment as soon as scheduling permitted. During the in-person pretreatment assessment, spouses completed questionnaires (including MCTS, Dominance/Isolation Scale, Fear of Spouse, Attribution of Responsibility Items, BDI, DAS, and SVPC) and interviews in separate rooms. Participants were not paid for their participation in the initial assessment or treatment phases of the study, but the treatment was free.

Couples who met the inclusion criteria were placed on the waiting list for the next available group. When 6 to 8 couples qualified, a new group was started. Groups alternated between Physical Aggressive Couples Treatment (PACT) and GST.<sup>2</sup> Therapists provided both types of treatment. Five groups of GST and seven groups of PACT were conducted. Forty couples were assigned to PACT; 22 (55%) completed ten out of fourteen 2-hour sessions and were considered treatment completers. Thirty couples were assigned to GST; 15 (50%) completed at least ten sessions.

Before each treatment session, spouses completed the CTS and answered additional questions about the level of disclosure in the previous week's ses-

<sup>2</sup> We alternated PACT and GST for groups 1 through 8. However, Dr. Neidig, the PACT supervisor, became seriously ill during the course of the study. In order to complete as many couples groups as possible before his eventual death, we stopped alternating groups and ran three couples groups in a row (groups 9 through 11), followed by a final GST group (group 12).

sion, whether anything discussed in the group led to physical aggression, and their attributions for the week's conflicts. Therapists spent the early part of each session asking each participant about the degree of conflict during the week.

At the end of the 14 week treatment, spouses completed the same questionnaire packet as at pretreatment (including MCTS, Dominance/Isolation Scale, Fear of Spouse, Attribution of Responsibility Items, BDI, DAS, and SVPC). In addition, the consumer satisfaction measure was included. Spouses were also interviewed privately by a research assistant and were given phone numbers for further treatment resources.

One year following the completion of treatment, spouses were again assessed with the questionnaire packet and a brief interview. Those who could not come to the clinic were mailed packets and completed a brief, confidential telephone interview. Each spouse was paid \$40 for completing the full follow-up assessment. If repeated attempts at scheduling appointments and mailing booklets failed, spouses were assessed by telephone only with the MCTS and the MAT (which was then converted to a DAS score). These spouses were paid \$20 each.

*Women's GST.* Women's groups comprised 6 to 8 women and a female therapist. The focus of the treatment was to help women: (a) recognize the characteristics of abusive relationships; (b) understand the emotional effects of violence; (c) learn ways to control their emotional reactions to negative events; and (d) evaluate the status of the marriage, and the advantages and disadvantages of staying in the marriage. A session-by-session summary can be found in Table 1. Therapists followed a detailed treatment manual (40 pages); women were given workbooks of approximately 100 pages.

*Men's GST.* Men's groups comprised 6 to 8 men and a male therapist. The focus of the treatment was to help men to: (a) decrease their use of psychological and physical aggression; (b) accept responsibility for their aggression; (c) understand the negative effects of violence on every member of the family; (d) learn to recognize the cycle of violence and to control their anger; and (e) learn how to communicate requests rather than give orders. A session-by-session summary can be found in Table 1. Therapists followed a detailed treatment manual (37 pages); men were given workbooks of approximately 100 pages.

*PACT.* PACT groups comprised 6 to 8 couples plus male and female co-therapists. The focus of the treatment was to help spouses: eliminate psychological and physical violence in the home; accept responsibility for escalation of angry interchanges and the resulting violence; recognize and control self-angering thoughts; communicate more effectively; increase caring and mutually pleasurable activities; and understand that each partner has a right to be treated with respect. A session-by-session summary can be found in Table 1. Therapists followed a detailed treatment manual (42 pages), supplemented with the Neidig and Friedman (1984) book; spouses were given workbooks of approximately 100 pages. Heyman and Neidig (1997) describe PACT in detail, including clinical transcripts.



TABLE 1  
SESSION-BY-SESSION SUMMARIES OF GROUP TREATMENTS FOR WIFE ABUSE

| Session | GST-Men                                                                          | GST-Women                                                                                    | PACT                                                                 |
|---------|----------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|----------------------------------------------------------------------|
| 1       | Introduction; recounting violent incident                                        | Introduction; telling your story; safety plan                                                | Introduction; recounting violent incident                            |
| 2       | Accepting responsibility for violence                                            | Characteristics of abusive relationships; impact of societal characteristics of spouse abuse | Walker's cycle of violence; discriminating different levels of anger |
| 3       | Walker's cycle of violence; time out                                             | Determining what you can change                                                              | Discriminating different levels of anger, time out procedures        |
| 4       | Cognitive behavioral model of anger                                              | Emotional effects of violence; self-esteem                                                   | Cognitive behavioral model of anger                                  |
| 5       | Anger control techniques; challenging hot thoughts                               | Cognitive behavioral model of depression; combating depression                               | Anger control techniques; challenging hot thoughts                   |
| 6       | Irrational thoughts                                                              | Emotional arousal: anger; responsibility for emotional arousal                               | Stress-abuse connection; irrational beliefs                          |
| 7       | Midterm progress evaluation; review                                              | Midterm progress evaluation; review                                                          | Midterm progress evaluation; review                                  |
| 8       | Family-of-origin lessons about anger; self-esteem                                | Human rights; assertiveness training                                                         | Communication principles and skills; positive behaviors              |
| 9       | Gender roles; identifying feelings                                               | Constructive communication; barriers to effective communication                              | Gender differences in communication; expressing feelings; empathy    |
| 10      | Power/control vs. equality tactics; characteristics of an equitable relationship | Increasing support resources; helping others help you                                        | Assertion vs. aggression; equality in rights and decision making     |
| 11      | Relaxation training                                                              | Dealing with criticism; taking constructive action                                           | Conflict escalation process; principles of conflict containment      |
| 12      | Assertiveness training                                                           | Power/control vs. equality tactics; evaluating your marriage                                 | Dirty fighting techniques                                            |
| 13      | Constructive communication; empathizing with others                              | Marriage as a choice                                                                         | Sex; jealousy; expanding social support network                      |
| 14      | Wrap-up; maintaining gains                                                       | Wrap-up; maintaining gains                                                                   | Wrap-up; maintaining gains; expressive versus instrumental violence  |

*Note.* GST: Gender Specific Treatment; PACT: Physical Aggression Couples Treatment.

The major theoretical difference between the two treatments is that in GST the male is held fully responsible for the physical aggression in the marriage, whereas in PACT, each partner is held responsible for the control of his or her own physical aggression. Furthermore, PACT encourages both partners to accept responsibility for their contribution to the process of conflict escalation. In both treatments, the greater impact of both physical and psychological aggression on women than men (Cantos et al., 1994; Cascardi, et al., 1992) is emphasized. In GST, a greater emphasis is placed on power, societal and patriarchal factors causing wife abuse than in the PACT.

*Adherence to treatment protocols.* Two bachelor-, and one graduate-level research assistant rated randomly selected treatment session videotapes for themes developed and subject matter covered by the therapists. Tapes were randomly assigned to coders; coders were blind as to which tapes were randomly chosen to be used as interrater reliability checks. Seventy-seven (i.e., 34% of the 225 total session videotapes) 2-hour sessions were coded (25 PACT, 22 GST-Men; 30 GST-Women); 20 sessions (i.e., 26% of the 77 coded tapes; 9 PACT, 5 GST-Men, 6 GST-Women) were coded by two raters to assess for interrater reliability. Differences in number of tapes coded were due to the random selection procedure.

Coders used a system developed for this study that assessed both the thematic content of the session (i.e., the overall conceptual thrust of therapist interventions) and the didactic content of the session (i.e., the overt material covered in the session). Three themes explicated PACT (e.g., conflict escalation is due to a mutual process involving circular causality) and three explicated GST (e.g., physical aggression is due to men's abuse of power and control tactics). Themes were rated on an occurrence/non-occurrence basis, and all had adequate interrater reliabilities (Cohen's  $\kappa$ ), as shown in Table 2. Sixteen content areas, comprising material from any of the three groups, were rated on an ordinal scale (i.e., on a 0 [no emphasis] to 3 [heavily emphasized] scale). Interrater reliabilities (Finn's  $r$ ), as shown in Table 3, were adequate, except for content areas of communication skills. Finn's  $r$  is a variant of the intra-class correlation and is preferable when ordinal data have even small degrees of skewness and kurtosis (Whitehurst, 1984).

*Therapists.* Three male and three female therapists conducted the groups. Of the six therapists, four had doctorates in clinical psychology, one had a degree in social work, and one was a clinical psychology doctoral candidate. Therapists received 2 days of training in each treatment by the supervisors. Because of logistic constraints (e.g., 2 therapists moved after the first year of the study on account of professional commitments), four pairs were used. Male therapist #1 and female therapist #1 conducted three groups (2 PACT, 1 GST) and then moved out of the state. Male therapist #2 and female therapist #2 conducted four groups (2 PACT, 2 GST). Male therapist #3 and female therapist #3 conducted three groups (2 PACT, 1 GST). Male therapist #3 and female therapist #2 conducted two groups (1 PACT, 1 GST). We attempted to assign an equal number of groups to each therapist; again, practical consider-

ation made actual assignments vary somewhat. Each therapist did, however, conduct at least one group in each treatment condition.

In GST, men had a male therapist and women had a female therapist. Thus, there was a match on sex. In PACT, a mixed sex co-therapy team conducted the mixed-sex groups. We did this to (a) provide equivalent client/therapist ratios in GST and PACT; (b) make sure that all spouses had a same sex therapist either leading or co-leading their group; and (c) match mode of therapy with sex of therapist (i.e., men's groups are led by a man, women's groups by a woman, and couple groups by a male/female team). As designed, GST and PACT groups differ on (a) the number of therapists; and (b) the presence of an opposite sex therapist.

*Supervisors.* The supervisors of the women's and men's groups were Dr. Ileana Arias and Dr. Alan Rosenbaum, respectively. The PACT groups were supervised by Dr. Neidig. Each treatment supervisor had conducted research and clinical work in family violence for more than 10 years. After each session, videotapes were duplicated and sent to supervisors. Supervisors watched the tapes prior to providing weekly supervision to maximize adherence to the treatments.

## Results

### *Adherence to Treatment Protocols*

Tables 2 and 3 indicate that the programs were administered as intended. For each of the six general themes of treatment (root causes or partner aggression, reasons for current conflict, necessary elements in change), the GST groups significantly differed from PACT in the expected directions. Specific content was somewhat less distinct, but still had significant differences in 7 of the sixteen areas (e.g., men's treatment emphasized male power and control more than did PACT; women's treatment emphasized marriage as a choice more than the other two groups did). Furthermore, the contents of the groups were never designed to be completely orthogonal, but representative of GST and PACT groups available elsewhere (e.g., many men's groups include some content on positive communication). However, the philosophical context in which the content was covered was designed to be quite distinct between GST and PACT, and Table 2 demonstrates that it was. In summary, even if the general topic were the same in the two groups (e.g., communication), the meaning was most likely distinct because the reasons for covering it were different (e.g., GST: Men changing their controlling behaviors; PACT: improve mutual communication skills).

### *Differential Effects of GST and PACT*

Means and standard deviations for mild and severe physical aggression in both GST and PACT at pretreatment, posttreatment, and follow-up are presented in Table 4. To test if decreases were significant, a  $(2 \times 2) \times 2$  repeated measures MANOVA was conducted, with mild and severe aggression as the

TABLE 2  
MEANS, STANDARD DEVIATIONS, AND INTER-RATER RELIABILITY FOR THERAPEUTIC THEMES  
OF WIFE ABUSE GROUPS

| Theme                                                                                                                                                  | GST-Men           |           | GST-Women         |           | PACT              |           | Reliability<br>$\kappa$ |
|--------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|-----------|-------------------|-----------|-------------------|-----------|-------------------------|
|                                                                                                                                                        | <i>M</i>          | <i>SD</i> | <i>M</i>          | <i>SD</i> | <i>M</i>          | <i>SD</i> |                         |
| 1. Root cause of partner aggression—natural couple disagreements                                                                                       | 0.05 <sub>a</sub> | 0.22      | 0.00 <sub>a</sub> | 0.00      | 0.59 <sub>b</sub> | 0.50      | .57                     |
| 2. Root cause of partner aggression—male prerogative, male control, societal support                                                                   | 0.24 <sub>a</sub> | 0.44      | 0.21 <sub>a</sub> | 0.41      | 0.00 <sub>b</sub> | 0.00      | .95                     |
| 3. Current conflict: Escalation due to mutual influence (i.e., circular causality), responsibility from both partners                                  | 0.00 <sub>a</sub> | 0.00      | 0.10 <sub>a</sub> | 0.31      | 0.93 <sub>b</sub> | 0.27      | 1.00                    |
| 4. Current conflict: Male alone is responsible for verbal and physical aggression                                                                      | 0.48 <sub>a</sub> | 0.51      | 0.36 <sub>a</sub> | 0.49      | 0.00 <sub>b</sub> | 0.00      | .73                     |
| 5. Change: Accepting conflict, but personal dealing with it better (Both spouses must change their ways to avoid escalating conflict)                  | 0.00 <sub>a</sub> | 0.00      | 0.07 <sub>a</sub> | 0.26      | 1.00 <sub>b</sub> | 0.00      | .90                     |
| 6. Change: Focus on self Men must change the attitudes and behaviors that promote male violence. Women should focus on their own safety and well-being | 1.00 <sub>a</sub> | 0.00      | 1.00 <sub>a</sub> | 0.00      | 0.07 <sub>b</sub> | 0.27      | .79                     |

Note. GST: Gender Specific Treatment; PACT: Physical Aggression Couples Treatment. Means with different subscripts differ significantly at the  $p < .05$  level by the Student-Newman-Keuls Significant Difference Test. Ratings scale was 0 = no, 1 = yes.

TABLE 3  
MEANS, STANDARD DEVIATIONS, AND INTER-RATER RELIABILITY FOR CONTENT  
OF WIFE ABUSE GROUPS

| Content Area                                                                                            | GST-Men             |           | GST-Women           |           | PACT              |           | Finn's<br><i>r</i> |
|---------------------------------------------------------------------------------------------------------|---------------------|-----------|---------------------|-----------|-------------------|-----------|--------------------|
|                                                                                                         | <i>M</i>            | <i>SD</i> | <i>M</i>            | <i>SD</i> | <i>M</i>          | <i>SD</i> |                    |
| Anger control—ABC model,<br>recognizing anger signals, cooling<br>off, considering consequences         | 1.38 <sub>a</sub>   | 1.12      | 0.55 <sub>b</sub>   | 0.99      | 1.22 <sub>a</sub> | 1.19      | 0.81               |
| Assertiveness—assertive vs.<br>nonassertive vs. aggressive<br>responses/behaviors                       | 0.62 <sub>a,b</sub> | 1.12      | 1.21 <sub>a</sub>   | 1.29      | 0.22 <sub>b</sub> | 0.64      | 0.98               |
| Communication skills                                                                                    | 0.71                | 1.01      | 0.76                | 1.09      | 1.26              | 1.38      | 0.37               |
| Conflict containment/time out                                                                           | 1.33 <sub>a</sub>   | 1.20      | 0.28 <sub>b</sub>   | 0.53      | 1.52 <sub>a</sub> | 1.16      | 0.83               |
| Definition of violence (types: physical,<br>psychological, sexual)                                      | 0.14                | 0.65      | 0.14                | 0.58      | 0.44              | 0.93      | 0.58               |
| Effects of violence on family, self,<br>partner; emotional reactions: fear,<br>depression, anxiety      | 0.19                | 0.68      | 0.31                | 0.85      | 0.15              | 0.46      | 0.58               |
| Empathy—encourage capacity for<br>empathy—consider spouses' point<br>of view and feelings               | 0.14                | 0.48      | 0.17                | 0.54      | 0.37              | 0.63      | 0.96               |
| Irrational beliefs—self-defeating/<br>irrational thoughts, techniques to<br>combat anger and depression | 0.19                | 0.60      | 0.24                | 0.58      | 0.00              | 0.00      | 0.98               |
| Jealousy—establishing boundaries,<br>rules to contain jealousy                                          | 0.14                | 0.36      | 0.00                | 0.00      | 0.22              | 0.58      | 0.92               |
| Marriage as a choice/personal change<br>consideration                                                   | 0.14 <sub>a</sub>   | 0.36      | 1.14 <sub>b</sub>   | 0.95      | 0.07 <sub>a</sub> | 0.27      | 0.87               |
| Personal responsibility "You alone are<br>responsible for your behavior."                               | 0.90 <sub>a</sub>   | 0.77      | 1.29 <sub>a,b</sub> | 1.05      | 0.56 <sub>b</sub> | 0.75      | 0.81               |
| Power and control (equitable vs.<br>inequitable tactics/relationships)                                  | 0.67 <sub>a</sub>   | 1.11      | 0.45 <sub>a,b</sub> | 0.95      | 0.07 <sub>b</sub> | 0.38      | 0.98               |
| Self esteem—interpreting events;<br>defeating vs. enhancing self-<br>statements                         | 0.52 <sub>a</sub>   | 1.08      | 0.76 <sub>a</sub>   | 1.06      | 0.00 <sub>b</sub> | 0.00      | 0.79               |
| Sex/marital/gender roles—messages<br>learned as child about male/female<br>roles                        | 0.52                | 0.93      | 0.31                | 0.66      | 0.19              | 0.62      | 1.00               |
| Social support/resources—evaluating<br>your support system, developing<br>resources & skills            | 0.00                | 0.00      | 0.17                | 0.60      | 0.19              | 0.62      | 0.98               |
| Stress management/relaxation<br>techniques—exercises to practice,<br>deep breathing, imagery            | 0.48                | 1.08      | 0.07                | 0.26      | 0.33              | 0.78      | 0.87               |

*Note.* GST: Gender Specific Treatment; PACT: Physical Aggression Couples Treatment. Means with different subscripts differ significantly at the  $p < .05$  level by the Student-Newman-Keuls Significant Difference Test. Ratings scale was 0 = no emphasis; 1 = rarely emphasized; 2 = somewhat emphasized; 3 = heavily emphasized.

TABLE 4  
PRE- AND POSTTREATMENT FREQUENCY OF PHYSICAL AGGRESSION ON THE CONFLICT TACTICS  
SCALE FOR TWO WIFE ABUSE TREATMENTS

|                            | Physical Aggression<br>Past 14 Weeks |           |               |           | Physical Aggression<br>Past Year |           |           |           |
|----------------------------|--------------------------------------|-----------|---------------|-----------|----------------------------------|-----------|-----------|-----------|
|                            | Pretreatment                         |           | Posttreatment |           | Pretreatment                     |           | Follow-up |           |
|                            | <i>M</i>                             | <i>SD</i> | <i>M</i>      | <i>SD</i> | <i>M</i>                         | <i>SD</i> | <i>M</i>  | <i>SD</i> |
| Husband-to-Wife Aggression |                                      |           |               |           |                                  |           |           |           |
| GST ( <i>n</i> = 14)       |                                      |           |               |           |                                  |           |           |           |
| Mild aggression            | 6.01                                 | 7.58      | 3.36          | 4.99      | 16.71                            | 12.34     | 9.64      | 12.83     |
| Severe aggression          | 1.12                                 | 1.61      | 0.64          | 1.08      | 2.50                             | 2.95      | 2.00      | 3.98      |
| PACT ( <i>n</i> = 23)      |                                      |           |               |           |                                  |           |           |           |
| Mild aggression            | 8.01                                 | 8.26      | 2.91          | 4.53      | 19.48                            | 15.99     | 7.95      | 11.96     |
| Severe aggression          | 2.45                                 | 3.52      | 0.95          | 2.26      | 3.65                             | 3.98      | 1.52      | 3.78      |
| Wife-to-Husband Aggression |                                      |           |               |           |                                  |           |           |           |
| GST ( <i>n</i> = 14)       |                                      |           |               |           |                                  |           |           |           |
| Mild aggression            | 6.07                                 | 8.38      | 2.36          | 4.36      | 12.71                            | 16.01     | 7.00      | 10.40     |
| Severe aggression          | 1.31                                 | 3.28      | 0.79          | 2.12      | 3.86                             | 6.76      | 2.00      | 4.00      |
| PACT ( <i>n</i> = 23)      |                                      |           |               |           |                                  |           |           |           |
| Mild aggression            | 5.83                                 | 5.51      | 1.86          | 2.27      | 16.00                            | 19.62     | 4.45      | 5.61      |
| Severe aggression          | 1.70                                 | 2.24      | 0.45          | 0.80      | 5.13                             | 8.84      | 0.96      | 1.44      |

Note. GST: Gender Specific Treatment; PACT: Physical Aggression Couples Treatment.  
PACT *n* = 22 at posttreatment and *n* = 21 at posttreatment due to missing data.

repeated measures, assessment (pretreatment vs. posttreatment) and aggressor (husband or wife) as the within-subject factors, and treatment type (GST or PACT) as the between-subject factor.<sup>3</sup> A similar analysis comparing pretreatment and follow-up<sup>4</sup> was also conducted.

Physical aggression decreased significantly from pretreatment to posttreatment, with no significant effect for gender of aggressor or treatment type. Subsequent univariate tests indicated that both mild aggression  $F(1, 34) = 13.34, p < .001$  and severe aggression  $F(1, 34) = 10.61, p < .005$  decreased significantly from pretreatment to posttreatment. Physical aggression also decreased significantly from pretreatment to follow-up, but with no significant effect for treatment type (see MANOVA in Table 5). Subsequent univariate tests indicated significant pre-follow-up decreases in both mild aggression  $F(1, 31) = 18.06, p < .001$  and severe aggression  $F(1, 31) = 6.57, p < .05$ . The significant main effect for aggressor was due to women having lower

<sup>3</sup> As an added measure, we replicated these analyses with covariance analyses, with little change in results.

<sup>4</sup> Two forms of the MCTS were administered at pretreatment. The pre-post comparison uses the MCTS with a 14-week time frame (to parallel the length of treatment), and the pre-follow-up uses the MCTS with a 1-year time frame (to parallel the length of follow-up).

TABLE 5  
PRETREATMENT AND POSTTREATMENT SCORES FOR TREATMENT COMPLETERS

| Scale <sup>a</sup>                                     | Pretreatment |           | Posttreatment |           | <i>t</i> | <i>df</i> |
|--------------------------------------------------------|--------------|-----------|---------------|-----------|----------|-----------|
|                                                        | <i>M</i>     | <i>SD</i> | <i>M</i>      | <i>SD</i> |          |           |
| Mild physical aggression (MCTS) during past 14 weeks   |              |           |               |           |          |           |
| Husband-to-wife                                        | 6.90         | 7.77      | 3.08          | 4.65      | 3.18***  | 35        |
| Wife-to-husband                                        | 5.81         | 6.68      | 2.06          | 3.20      | 3.91**** | 35        |
| Severe physical aggression (MCTS) during past 14 weeks |              |           |               |           |          |           |
| Husband-to-wife                                        | 1.70         | 2.61      | 0.83          | 1.87      | 2.88***  | 35        |
| Wife-to-husband                                        | 1.51         | 2.67      | 0.58          | 1.44      | 2.99***  | 35        |
| Dominance/isolation                                    |              |           |               |           |          |           |
| Husband-to-wife                                        | 27.60        | 11.47     | 20.89         | 8.14      | 5.05**** | 34        |
| Wife-to-husband                                        | 24.78        | 7.07      | 20.38         | 5.63      | 4.27**** | 31        |
| Psychological aggression (MCTS) during past 14 weeks   |              |           |               |           |          |           |
| Husband-to-wife                                        | 62.84        | 35.12     | 33.08         | 27.30     | 4.19**** | 35        |
| Wife-to-husband                                        | 37.02        | 29.59     | 30.36         | 30.32     | 1.64     | 35        |
| Fear                                                   |              |           |               |           |          |           |
| Husband's fear of wife                                 | 6.38         | 3.09      | 4.56          | 4.17      | 2.82***  | 31        |
| Wife's fear of husband                                 | 7.57         | 3.36      | 5.80          | 4.26      | 2.56**   | 34        |
| Beck Depression Inventory                              |              |           |               |           |          |           |
| Husband                                                | 8.68         | 6.28      | 6.26          | 6.47      | 2.48**   | 30        |
| Wife                                                   | 13.15        | 6.95      | 7.44          | 7.55      | 5.61**** | 34        |
| Attributions for aggression (Men)                      |              |           |               |           |          |           |
| H→W aggression caused by husband                       | 3.71         | 1.15      | 3.07          | 1.44      | 2.93**   | 27        |
| H→W aggression caused by wife                          | 1.73         | 0.93      | 2.62          | 1.42      | 3.23**   | 25        |
| Attributions for aggression (Women)                    |              |           |               |           |          |           |
| H→W aggression caused by husband                       | 2.72         | 1.39      | 2.66          | 1.23      | 0.29     | 28        |
| H→W aggression caused by wife                          | 2.31         | 1.44      | 2.72          | 1.39      | 1.33     | 28        |
| Dyadic Adjustment Scale                                |              |           |               |           |          |           |
| Husbands                                               | 83.48        | 14.63     | 93.71         | 15.39     | 4.36**** | 30        |
| Wives                                                  | 74.95        | 18.66     | 87.72         | 16.92     | 5.20**** | 34        |
| Spouse Verbal Problem Checklist                        |              |           |               |           |          |           |
| Husbands' problems (wives' report)                     | 42.75        | 10.92     | 41.52         | 11.87     | .61      | 34        |
| Wives' problems (husbands' report)                     | 39.35        | 11.16     | 40.86         | 12.98     | .87      | 31        |

*Note.* MCTS: Modified Conflict Tactics Scale (14-week time frame); Dominance/Isolation from Tolman Psychological Maltreatment of Women scale.

<sup>a</sup> *ns* differ due to missing data.

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .005$ ; \*\*\*\*  $p < .001$ .

scores than men when pretreatment and follow-up scores were collapsed. In short, men's and women's physical aggression decreased significantly, both at posttreatment and follow-up. However, neither treatment type significantly outperformed the other.

We also tested for differential effects of the treatment types on dominance/

isolation, psychological aggression, fear, depressive symptomatology, attributions for aggression, marital adjustment and communication problems. Each dependent variable was tested in 2 (pre-post)  $\times$  2 (treatment type) and 2 (pre-follow-up)  $\times$  2 (treatment type) repeated measures MANOVAs. There were no differential effects of treatment, except for the hypothesized finding that PACT husbands' marital adjustment at posttreatment improved more than GST husbands' did (prepost  $\times$  treatment type: Wilks'  $\lambda = .82$ ,  $F(1, 27) = 6.44$ ,  $p < .05$ ; pretreatment, GST:  $M = 87.20$ ,  $SD = 12.43$ , PACT:  $M = 81.71$ ,  $SD = 15.54$ ; posttreatment, GST:  $M = 89.49$ ,  $SD = 9.11$ , PACT  $M = 95.71$ ,  $SD = 17.45$ ).

We had hypothesized greater improvement for women's marital adjustment in PACT than in GST, and were surprised that both improved significantly. Thus, we were curious to see whether different processes were associated with wives' improved marital adjustment in PACT and GST. As such, a 2  $\times$  2 repeated measures MANOVA was conducted, with wives' DAS as the repeated measure, pre-post as the within-groups factor, format (GST or PACT) as the independent variable, and husbands' pre-post changes in mild and severe physical aggression, MCTS psychological abuse, dominance/isolation behaviors, and communication problems as the covariates. Differential change by format was found for the association between wives' DAS change and husbands' communication problems:  $F(2, 24) = 6.52$ ,  $p < .005$  and wives' DAS change and husbands' dominance/isolation behaviors:  $F(2, 24) = 3.54$ ,  $p < .05$ . Subsequent univariate analyses indicated that, in PACT, only communication problems were associated with wives' DAS change:  $F(1, 15) = 12.19$ ,  $p < .005$ , whereas in GST, only dominance/isolation behaviors approached statistically significant association with wives' DAS change:  $F(1, 8) = 3.50$ ,  $p < .10$ . Thus, improvements in wives' marital adjustment fit with the respective targets of intervention in the two groups. Due to missing data, sample sizes were too small to replicate the analyses at follow-up.

### *Overall Effects of Treatments*

Because there were few significant differential treatment effects, the analyses below combine completers of both treatment types to examine overall effects of treatment.<sup>5</sup> Table 5 contains means, standard deviations, and test statistics for all pre-post comparisons. Table 6 contains similar information for all pre-follow-up comparisons. Findings from these tables will be summarized below.

*Dominance/isolation and psychological aggression.* Husbands' dominance/isolation behaviors and husband-to-wife psychological aggression (on the MCTS) dropped significantly between the 14-week pretreatment period and the 14 weeks during treatment. At posttreatment, wives' dominance/isolation behaviors decreased significantly, but not wife-to-husband psychological aggression. At follow-up, husbands' and wives' psychological aggression

<sup>5</sup> Full intercorrelation matrices are available upon request on the authors' web page at <http://www.psy.sunysb/marital/>



TABLE 6  
PRETREATMENT AND FOLLOW-UP SCORES FOR TREATMENT COMPLETERS

| Scale <sup>a</sup>                                 | Pretreatment |           | Follow-Up |           | <i>t</i> | <i>df</i> |
|----------------------------------------------------|--------------|-----------|-----------|-----------|----------|-----------|
|                                                    | <i>M</i>     | <i>SD</i> | <i>M</i>  | <i>SD</i> |          |           |
| Mild physical aggression (MCTS) during past year   |              |           |           |           |          |           |
| Husband-to-wife                                    | 19.31        | 14.51     | 8.63      | 12.15     | 3.68**** | 34        |
| Wife-to-husband                                    | 14.97        | 18.81     | 5.50      | 7.90      | 3.16***  | 33        |
| Severe physical aggression (MCTS) during past year |              |           |           |           |          |           |
| Husband-to-wife                                    | 3.34         | 3.69      | 1.71      | 3.81      | 2.07*    | 34        |
| Wife-to-husband                                    | 4.76         | 8.38      | 1.39      | 2.79      | 2.37*    | 33        |
| Dominance/isolation                                |              |           |           |           |          |           |
| Husband-to-wife                                    | 24.43        | 6.82      | 22.90     | 6.60      | 1.38     | 20        |
| Wife-to-husband                                    | 28.00        | 10.79     | 25.76     | 10.38     | 1.20     | 24        |
| Psychological aggression (MCTS) during past year   |              |           |           |           |          |           |
| Husband-to-wife                                    | 93.37        | 45.17     | 62.88     | 44.79     | 2.93***  | 34        |
| Wife-to-husband                                    | 76.64        | 44.46     | 44.06     | 34.20     | 5.49**** | 33        |
| Fear                                               |              |           |           |           |          |           |
| Husband's fear of wife                             | 6.30         | 3.15      | 4.59      | 3.26      | 2.54**   | 26        |
| Wife's fear of husband                             | 7.72         | 3.42      | 7.97      | 4.13      | 0.35     | 28        |
| Beck Depression Inventory                          |              |           |           |           |          |           |
| Husband                                            | 8.22         | 5.75      | 6.77      | 5.64      | 1.08     | 26        |
| Wife                                               | 12.39        | 5.60      | 8.79      | 7.81      | 2.78***  | 28        |
| Attributions for aggression (Men)                  |              |           |           |           |          |           |
| H→W aggression caused by husband                   | 3.44         | 1.32      | 2.75      | 1.39      | 2.30*    | 14        |
| H→W aggression caused by wife                      | 1.80         | 1.08      | 2.27      | 1.54      | 1.33     | 14        |
| Attributions for aggression (Women)                |              |           |           |           |          |           |
| H→W aggression caused by husband                   | 2.68         | 1.46      | 2.86      | 1.17      | 0.77     | 21        |
| H→W aggression caused by wife                      | 2.19         | 1.29      | 2.76      | 1.39      | 1.55     | 20        |
| Dyadic Adjustment Scale <sup>b</sup>               |              |           |           |           |          |           |
| Husbands                                           | 81.60        | 13.46     | 93.02     | 14.05     | 3.26***  | 24        |
| Wives                                              | 73.29        | 19.53     | 82.98     | 17.59     | 2.38*    | 25        |
| Spouse Verbal Problem Checklist                    |              |           |           |           |          |           |
| Husband's problems (wife's report)                 | 41.96        | 11.84     | 37.17     | 12.24     | 2.43*    | 26        |
| Wife's problems (husband's report)                 | 39.63        | 11.20     | 30.37     | 12.67     | 4.05**** | 28        |

*Note.* MCTS: Modified Conflict Tactics Scale (1-year time frame).

<sup>a</sup> *ns* differ due to missing data.

<sup>b</sup> Dyadic Adjustment Scale administered only to those who were still married at follow-up.

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .005$ ; \*\*\*\*  $p < .001$ .

decreased significantly from 1 year prior to treatment; dominance/isolation behaviors did not change significantly.

*Depressive symptomatology and fear.* Wives' depressive symptomatology (on the BDI) was significantly lower at posttreatment and follow-up than at

pretreatment. In addition, wives' fear of their husbands also decreased significantly from pretreatment to posttreatment. Their fear at follow-up, however, did not significantly differ from that at pretreatment. However, it should be kept in mind that we excluded couples where the wife feared participating in a group with her husband, and, in general, fear levels were relatively low.

*Beliefs about responsibility for husband-to-wife aggression.* At posttreatment and at follow-up, aggressive husbands reported significant increases in taking responsibility for their own aggression. At posttreatment, husbands reported significant decreases in placing responsibility for their own aggression on their wives. Wives did not significantly change (at posttreatment or follow-up) in their attributions of responsibility for their husbands' aggression. Counter to concerns that conjoint treatment may lead wives to blame themselves for husband-to-wife aggression, such attributions did not change differentially by treatment type. In brief, as predicted, husbands' attributions changed in the clinically desired direction (showed increased responsibility for their aggression), but wives' attributions did not change.

*Marital adjustment and communication problems.* Treatment led to increased marital adjustment, at both posttreatment and follow-up, for both men and women. Neither genders' communication problems (on the SVPC) improved by posttreatment, but both had improved significantly by the follow-up.

### *Clinically Significant Change*

*Cessation of physical aggression.* A statistically significant number of husbands ceased physical aggression during treatment. Whereas only 3 of 36 (8%) husbands were not aggressive in the 14 weeks prior to treatment, 14 of 36 (39%) were not aggressive during the 14 weeks of treatment,  $\chi^2$  for correlated proportions (1) = 8.07,  $p < .01$ . Similarly, 15 of 36 (42%) husbands were not severely aggressive in the 14 weeks prior to treatment, but during the 14 weeks of treatment, 23 of 36 (64%) were not severely aggressive,  $\chi^2$  for correlated proportions (1) = 3.56,  $p < .07$ . Thus, although 22 husbands continued to be aggressive, 13 of them severely aggressive, a large minority (39%) completely ceased aggression during treatment. Seven of the 13 severely aggressive husbands were in the GST condition and six were in the PACT condition.

All husbands were aggressive in the year prior to treatment (as required by the screening criteria). Nine of 35 (26%) husbands were not aggressive in the year following treatment,  $\chi^2$  for correlated proportions (1) = 9.00,  $p < .01$ . Whereas 8 of 35 (23%) husbands were not severely aggressive in the year prior to treatment, 23 of 35 (66%) were not severely aggressive in the year following treatment,  $\chi^2$  for correlated proportions (1) = 11.84,  $p < .001$ . Of the nine husbands who were violence free at the 1-year follow-up, five were non-aggressive during the treatment, three had one act of mild aggression, and one had two acts of mild aggression.

Similar cessation results were found for women. Whereas 29% were not aggressive in the 14 weeks prior to treatment, 47% were not aggressive dur-

ing the 14 weeks of treatment,  $\chi^2$  for correlated proportions (1) = 3.56,  $p < .07$ . Similarly, 42% were not severely aggressive in the 14 weeks prior treatment, but during the 14 weeks of treatment, 69% were not severely aggressive,  $\chi^2$  for correlated proportions (1) = 5.33,  $p < .05$ . Fourteen percent of wives were not aggressive in the year prior to treatment; 23% were not aggressive in the year following treatment,  $\chi^2$  for correlated proportions (1) = 1.29, ns. Whereas 34% were not severely aggressive in the year prior to treatment, 57% were not severely aggressive in the year following treatment,  $\chi^2$  for correlated proportions (1) = 4.57,  $p < .05$ .

*Marital distress.* We also calculated clinically significant change for marital distress (i.e., the number of spouses who moved from the distressed into the nondistressed range on the DAS; Jacobson, Follette, & Revenstorf, 1984; Jacobson & Truax, 1991). As suggested by Jacobson and Truax, we calculated the reliable change index ( $RC = x_2 - x_1 / S_{diff}$ ). If a participant improved more than one standard deviation on the reliable change index (i.e.,  $1.96 * RC$ ), he or she was considered to have improved. To be considered "recovered," participants had to improve beyond the confidence interval (i.e.,  $1.96 * S_{diff}$ ) above the clinical cutoff of 97 (Jacobson & Truax).<sup>6</sup> At post-treatment, clinically significant change was as follows—husbands: 21.9% recovered; 31.3% improved; 43.8% not improved; wives: 19.4% recovered; 38.9% improved; 36.1% not improved. At follow-up, clinically significant change was as follows—husbands: 23.1% recovered; 34.6% improved; 42.3% not improved; wives: 7.1% recovered; 57.1% not improved.

### *Safety of Women During Treatment*

The hypothesis of some feminists that conjoint treatment increases wives' risk of victimization more than gender specific treatment was not supported. Couples' arguments regarding content discussed in the treatment or something related to the treatment led to physical aggression in approximately 2% of the sessions for each of the groups. Likewise, the belief that wives would be more afraid to express themselves in PACT than in GST groups was not supported. Women in both groups reported (on a scale of 1 [not at all afraid] to 5 [afraid all the time]) little fear of expressing themselves during sessions (PACT:  $M = 1.57$ ,  $SD = .62$ ; GST:  $M = 1.47$ ,  $SD = .64$ ,  $t(31) = .41$ , ns).

### *Consumer Satisfaction*

Spouses were highly satisfied with both forms of treatment; there were no differences in ratings of consumer satisfaction between GST and PACT. Ratings on specific items (with 9 being extremely satisfied) were as follows: "Your interest and involvement in the program" (men:  $M = 8.34$ ,  $SD = .65$ ; women:  $M = 8.37$ ,  $SD = .77$ ); "How relevant the program was for you" (men:  $M = 8.06$ ,  $SD = .98$ ; women:  $M = 8.23$ ,  $SD = .97$ ); "The personality

<sup>6</sup> One unchanged husband was already in the reliably recovered range at pretreatment, and was excluded from the clinically significant change analyses.

of the therapist (men:  $M = 8.47$ ,  $SD = .67$ ; women:  $M = 8.77$ ,  $SD = .49$ ); "Therapist skill and competence" (men:  $M = 8.53$ ,  $SD = .62$ ; women:  $M = 8.79$ ,  $SD = .48$ ); "How much your own concerns/goals were met by the program" (men:  $M = 7.47$ ,  $SD = 1.72$ ; women:  $M = 7.77$ ,  $SD = 1.46$ ).

### *Treatment and Marital Separation during Follow-up Period*

Seventeen of the 37 couples reported that at least one spouse received treatment (e.g. individual, couple) during the follow-up. Wives' treatment seeking was significantly related to husbands' continued use of severe aggression; (7 of the 9 [78%] of severely victimized wives sought treatment, whereas 6 of the 19 [32%] of non-severely victimized wives sought treatment;  $\chi^2(1) = 5.24$ ,  $p < .05$ ). Eighty-nine percent of treatment seeking wives said that participation in our program made their subsequent treatment-seeking easier. Finally, two couples (one GST and one PACT) separated during the follow-up.

## Discussion

Much controversy exists about the appropriateness of conjoint treatment of wife abuse. With a sample of volunteer, intact couples, we found that commonly expressed fears about conjoint treatment by some who work with court mandated men (e.g., Adams, 1988; McMahon & Pence, 1996) did not apply to our setting. For example, compared to wives in GST, wives in the conjoint treatment were not fearful of participating with their husbands; were not fearful during the sessions; did not blame themselves for the violence; and were not put at an increased risk for violence during the program.<sup>7</sup> In addition, husbands in both treatment groups took significantly more responsibility for their aggression. These results imply that the primary message received from both programs was that husbands were solely responsible for their aggression. Furthermore, the adherence data indicated that the groups were quite distinct in their philosophies. We should emphasize that our conjoint treatment, PACT, was explicitly for husband-to-wife aggression, not standard couples therapy; our couples were carefully screened to ensure that, although all husbands were aggressive, wives had not been seriously injured, nor were they fearful of participating with their husbands; and both spouses wanted to participate and were willing to be assigned to either GST or conjoint groups. Like others, we strongly believe that it is important to delineate physically aggressive men in terms of severity, types, and/or impact of the aggression (Hamberger & Hastings, 1993; Holtzworth-Munroe & Stuart, 1994; O'Leary, 1993; Vivian & Malone, 1997).

The results can be viewed as both encouraging and dismaying. On the

<sup>7</sup> The weekly violence measure asked if "Anything discussed in last week's treatment session led to a *serious* physical argument" (emphasis added). Even though GST and PACT defined any act of physical aggression as serious, the wording of the question was problematic and may have underestimated the risk of aggression following treatment sessions across all types of treatment.

encouraging side, both programs resulted in equivalent improvements, at both posttreatment and 1 year follow-up, on a host of outcome measures, ranging from aggressive acts to communication problems. As hypothesized, participants in both forms of 14-week group treatment for wife abuse reported significant reductions, at both posttreatment and follow-up, in mild and severe aggression (both husband-to-wife and wife-to-husband); husbands' psychological aggression; wives' depressive symptomatology; husbands' taking responsibility for their aggression; and husbands' and wives' marital adjustment. Furthermore, significant pre-follow-up improvements in wives' psychological aggression and both partners' communication problems were found. Finally, consumer satisfaction ratings indicated that spouses who completed the program found it to have been helpful.

On the dismaying side, we fell far short of the primary goal of both the GST and PACT programs—to eliminate physically aggressive behavior in the home. Furthermore, during the year following treatment, three-quarters of the husbands were physically aggressive, with one third committing acts of severe aggression. Because three fourths of the husbands were severely aggressive in the year prior to treatment, the programs seem to have succeeded in the tertiary prevention of severe aggression—perhaps by inducing husbands to take self-corrective actions when they committed acts of mild aggression, instead of escalating their violence. Thus, one interpretation of this treatment effort was that physical aggression is very hard to change. That interpretation makes perfect sense and fits with the concept that physical aggression is one of the most stable characteristics of mankind (Olweus, 1979).

As might be expected, longitudinal evaluations of men who engage in moderate to severe aggressive behavior are sparse. A study by Jacobson, Gottman, Gortner, Berns, & Shortt (1996) assessed 26 aggressive men who remained with their partners across a 2-year period. This sample comprised couples in which the wives reported an average of approximately 21 acts of physical aggression by their husband in the past year, and the wives had an average marital satisfaction score (DAS) of approximately 94. Only 1 of the 26 men (4%) in the Jacobson et al. study ceased being physically aggressive. In this study, wives reported that their husbands engaged in approximately 22 acts of physical aggression against them in the past year; they had an average marital satisfaction score (DAS) of approximately 75. Thus, the two samples were approximately equivalent in terms of overall aggression as reported by the wives, but wives from our sample were much more discordant at pretreatment. In this study, wives reported that 33% of the husbands ceased being physically aggressive when contrasted with 4% of those in the Jacobson et al. study. Perhaps more striking, for men not in treatment in the Jacobson et al. study, psychological abuse did not decrease across time even when physical abuse did. In contrast, in this study, there was a 47% reduction in psychological abuse (that occurred in conjunction with approximately 50% reductions in both mild and severe physical aggression). This reduction in psychological aggression is important since it is becoming apparent that psychological abuse

is a better predictor of depression, marital discord, and divorce than is physical abuse (O'Leary, 1999).

Predictions of differential effects of the two treatment approaches found little support, despite evidence that the themes of each program were delivered appropriately. We had hypothesized that PACT would produce significant improvements on both aggression and marital adjustment, whereas GST would produce improvements on aggression but not marital adjustment because GST had little overt focus on the relationship skills or issues. This hypothesis held true only for husbands. At posttreatment, wives in both groups, and husbands in PACT, improved significantly on marital adjustment. Follow-up analyses revealed that, in accord with the content emphasized in the two treatments, improvements in marital adjustment for GST women were associated with their husbands' reduced use of dominance/isolation tactics, whereas for PACT women, improvements were associated with their husbands' reduction in problematic communication behaviors.

Despite the overall positive outcomes reported, two problem areas should be emphasized. First, the drop out rate in this study was 47%. This rate is similar to that often reported in the batterers' treatment literature (Hamberger & Hastings, 1993), and it is a significant problem for partner-abuse treatment programs nationally. There was a 15% dropout rate in the Brannen and Rubin (1996) study with men mandated to treatment. This dropout rate is lower than rates usually reported, and it was lower than the dropout rate in this study with men who volunteered for treatment. In one of the largest studies to date comparing men mandated to treatment and volunteers for treatment, Rosenbaum et al. (1997) found that volunteers had significantly higher dropout rates than those who were mandated. We assessed all dropouts and found that the majority of our couples said that they dropped out of treatment because a program focusing on aggression did not address enough of their own specific marital problems (Brown et al., 1997).

Second, our cessation rate of 26% at follow-up may be lower than that found in programs with court-mandated men. Edleson and Tolman's (1992) review of the literature reported rates of 53% to 85%, in follow-up periods ranging from a couple of weeks to 1 to 2 years. However, there are several issues that must be considered when comparing these rates. First, the treatment populations differ (court-mandated vs. volunteer). Second, men in court-mandated programs in various states frequently are no longer in contact with the victim (or have protective orders prohibiting contact), and thus may be more likely to appear to have ceased aggression. Third, involvement in legal proceedings may increase the likelihood of cessation. Fourth, in our study, 84% of the wives who completed treatment participated in the follow-up, whereas the follow-up participation of the completers is sometimes as low as 20% to 40% at 6–12 months (Gondolf, 1997).

The couples in this study responded to a notice of treatment for spouses whose "arguments led to throwing, pushing, and shoving." Some might argue that these couples are not like couples usually seeking marital therapy. How-

ever, one-half to two-thirds of couples seeking marital therapy are physically aggressive (Holtzworth-Munroe et al., 1992; O'Leary et al., 1992; Vivian & Langhinrichsen, 1994). They have arguments that lead to throwing, pushing, and shoving. Moreover, the couples herein reported problems remarkably similar to those couples presenting for routine marital therapy (O'Leary et al.). The most frequently reported problems of those participating were commitment (43%), communication (42%), and sexuality (25%). Finally, a mental health clinic or a marital clinic might well advertise the availability of treatment for couples whose arguments sometimes lead to aggression, just as clinics sometimes offer help to couples with communication or sexual problems.

Our therapists received training in both treatment modalities, they received weekly consultation by the experts in GST and PACT, and adherence ratings indicated that there were differences in therapeutic content across the two treatments. Prior to participating in the study, five of the six therapists had considerable experience both with individual and marital therapy. All therapists reported feeling comfortable delivering both treatments, and there were no differences in the effectiveness of the therapists. Research by the first author has been more on the role of marital variables as risk factors for abuse (Pan et al., 1994a) than on power and control variables, though many of our research projects, including this one, include measurement of both variables. Consequently, the results need to be interpreted in this light along with the prediction of less relapse at follow-up for those in PACT. Nonetheless, no differential changes were found in the Brannen and Rubin (1996) study or in the current study.

Several future directions are indicated. First, further work must be done to isolate the mediators and moderators of improvement in each type of treatment. Second, like the majority of interventions for wife abusers, our drop out rate needs to be reduced. Volunteer samples, especially those of intact couples who believe that aggression is secondary to other relationship problems (cf. Ehrensaft & Vivian, 1996), present challenges different from that found in court-mandated samples. Many couples from our study (Brown et al., 1997) said that they dropped out because they wanted individualized (as opposed to group) treatment. Thus, we believe that, for couples who do not identify aggression as a presenting problem, individualized programs should be tested that meet the couples' expressed interest in marital improvement as well as the therapist's interest in aggression abatement. Finally, further research is necessary to determine which treatment is optimal for which types of clients.

In conclusion, findings from our study and those of Brannen and Rubin (1996) demonstrate significant reductions in psychological and physical aggression for both GST and conjoint treatment with no differential effects. Brannen and Rubin served a court-mandated population, and we served a volunteer population. At this point, both conjoint treatment and GSTs for wife abuse appear to be equally viable modes of intervention. There were significant reductions in men's psychological aggression (i.e., approximately a 50% reduction) that were associated with reductions in physical aggression. Relat-

edly, there were significant increases in marital satisfaction for both men and women; 58% of the wives and 53% of the husbands showed reliable change (i.e., improvement or recovery).

Despite these marked reductions, the fact that physical aggression of men and women was reported at follow-up is of concern, and future research should address issues such as what predicts continued aggression or cessation and the meaning of the maintenance of physical aggression to the clients when psychological aggression is reduced and marital satisfaction is improved.

## References

- Adams, D. (1988). Treatment models of men who batter: A profeminist analysis. In K. Yllö & M. Bograd (Eds.), *Feminist perspectives on wife abuse*. (pp. 176–199). Newbury Park, CA: Sage.
- American Psychological Association. (1996). *Violence and the family: Report of the American Psychological Association Presidential Task Force on Violence and the Family*. Washington, DC: Author.
- Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8, 77–100.
- Brannen, S. J., & Rubin, A. (1996). Comparing the effectiveness of gender-specific and couples groups in a court-mandated-spouse-abuse treatment program. *Research on Social Work Practice*, 6, 405–424.
- Brown, P. D., O'Leary, K. D., & Feldbau, S. R. (1997). Dropout in a treatment program for self-referring wife abusing men. *Journal of Family Violence*, 12, 365–387.
- Caesar, P. L., & Hamberger, L. K. (Eds.). (1989). *Treating men who batter: Theory, practice, and programs*. New York: Springer.
- Cantos, A. L., Neidig, P. H., & O'Leary, K. D. (1994). Injuries of women and men in a treatment program for domestic violence. *Journal of Family Violence*, 9, 113–124.
- Cascardi, M., Langhinrichsen, J., & Vivian, D. (1992). Marital aggression, impact, injury, and health correlates for husbands and wives. *Archives of Internal Medicine*, 152, 1178–1184.
- Cascardi, M., & Vivian, D. (1995). Context for specific episodes of marital violence: Gender and severity of violence differences. *Journal of Family Violence*, 10, 265–293.
- Crane, D. R., Allgood, S. M., Larson, J. H., & Griffin, W. (1990). Assessing marital quality with distressed and nondistressed couples: A comparison and equivalency table for three frequently used measures. *Journal of Marriage and the Family*, 52, 87–93.
- Crowell, N. A., & Burgess, A. (1996). *Understanding violence against women*. Washington, DC: National Academy Press.
- Edleson, J. L., & Tolman, R. M. (1992). *Intervention for men who batter: An ecological approach*. Newbury Park, CA: Sage.
- Ehrensaft, M. K., & Vivian, D. (1996). Spouses reasons for not reporting existing marital aggression as a marital problem. *Journal of Family Psychology*, 10, 443–453.
- Fagan, J. (1996). *The criminalization of domestic violence: Promises and limits*. Washington, DC: National Institute of Justice Research.
- Ganley, A. L. (1981). *Court-mandated counseling for men who batter*. (Manual available from the Center for Women's Policy Studies, Washington, DC).
- Geffner, R., Mantooth, C., Franks, D., & Rao, L. (1989). A psychoeducational, conjoint therapy approach to reducing family violence. In P. L. Caesar & L. K. Hamberger (Eds.), *Treating men who batter: Theory, practice, and programs* (pp. 103–133). New York: Springer.
- Gondolf, E. W. (1997). Batterer programs: What we know and need to know. *Journal of Interpersonal Violence*, 12, 83–98.



- Hamberger, L. K. (1996). Group treatment of women who batter their female partners. *In Session*, 2, 25–48.
- Hamberger, L. K., & Hastings, J. E. (1988). Skills training for treatment of spouse abusers: An outcome study. *Journal of Family Violence*, 3(2), 121–130.
- Hamberger, L. K., & Hastings, J. E. (1993). Court-mandated treatment of men who batter their partners: Issues, controversies, and outcomes. In Z. Hilton (Ed.), *Legal responses to assault* (pp. 188–229). Newbury Park, CA: Sage.
- Hamberger, L. K., & Lohr, J. M. (1989). Proximal cause of spouse abuse: A theoretical analysis for cognitive behavioral interventions. In P. L. Caesar & L. K. Hamberger (Eds.), *Treating men who batter: Theory, practice, and programs* (pp. 53–100). New York: Springer.
- Haynes, S. N., Chavez, R. E., & Samuel, V. (1984). Assessment of marital communication and distress. *Behavioral Assessment*, 6, 315–322.
- Heyman, R. E., & Neidig, P. H. (1997). Physical aggression couples treatment. In W. K. Halford & H. J. Markman (Eds.), *Clinical handbook of marriage and couples intervention* (pp. 589–617). New York: Wiley.
- Heyman, R. E., Sayers, S. L., & Bellack, A. S. (1994). Global marital satisfaction vs. marital adjustment: Construct validity and psychometric properties of three measures. *Journal of Family Psychology*, 8, 432–446.
- Heyman, R. E., & Schlee, K. A. (1997). Toward a better estimate of the prevalence of partner abuse: Adjusting rates based on the sensitivity of the conflict tactics scale. *Journal of Family Psychology*, 11, 332–338.
- Heyman, R. E., Weiss, R. L., & Eddy, J. M. (1995). Marital interaction coding system: Revision and empirical evaluation. *Behavioural Research and Therapy*, 33, 737–746.
- Holtzworth-Munroe, A., & Stuart, G. L. (1994). Typologies of male batterers: Three subtypes and the differences among them. *Psychological Bulletin*, 116, 476–497.
- Holtzworth-Munroe, A., Waltz, J., Jacobson, N. S., Monaco, V., Fehrenbach, P. A., & Gottman, J. M. (1992). Recruiting non-violent men as control subjects for research on marital violence: How easily can it be done? *Violence and Victims*, 7, 79–88.
- Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, 15, 336–352.
- Jacobson, N. S., Gottman, J. M., Gortner, E., Berns, S., & Shortt, J. W. (1996). Psychological factors in the longitudinal course of battering: When do the couples split up? When does the abuse decrease? *Violence and Victims*, 11, 371–392.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12–19.
- Jouriles, E. N., & O'Leary, K. D. (1985). Interspousal reliability of reports of marital violence. *Journal of Consulting and Clinical Psychology*, 53, 419–421.
- Krokoff, L. J. (1989). Predictive validation of a telephone version of the Locke-Wallace Marital Adjustment Test. *Journal of Marriage and the Family*, 51, 767–775.
- Lindquist, C. U., Telch, C. F., & Taylor, J. (1983). Evaluation of a conjugal violence treatment program: A pilot study. *Behavioral Counseling and Community Intervention*, 3, 76–90.
- Locke, H. J., & Wallace, K. M. (1959). Short marital adjustment and prediction tests: Their reliability and validity. *Marriage and Family Living*, 21, 251–255.
- Malone, J., & O'Leary, K. D. (1995, April). *Continuation and cessation of violence in young married couples*. Paper presented at the Southern Sociological Society, Atlanta, GA.
- McMahon, M., & Pence, E. (1996). Replying to Dan O'Leary. *Journal of Interpersonal Violence*, 11, 452–455.
- Moffitt, T. E., Caspi, A., Margolin, G., Kreuger, Magdol, L., Silva, P. A., & Sydney, R. (1997). Do partners agree about abuse in their relationship? A psychometric evaluation of inter-partner agreement. *Psychological Assessment*, 9, 47–56.

- Neidig, P. H., & Friedman, D. (1984). *Spouse abuse: A treatment program for couples*. Champaign, IL: Research Press.
- O'Leary, K. D. (1993). Through a psychological lens: Personality traits, personality disorders, and levels of violence. In R. J. Gelles & D. R. Loseke (Eds.), *Current controversies on family violence* (pp. 7-30). Newbury Park, CA: Sage.
- O'Leary, K. D. (1996). Physical aggression in intimate relationships can be treated within a marital context under certain circumstances. *Journal of Interpersonal Violence*, 11, 450-452.
- O'Leary, K. D. (1999). Psychological abuse: A variable deserving critical attention in domestic violence. *Violence and Victims*, 14, 25-35.
- O'Leary, K. D., Barling, J., Arias, I., Rosenbaum, A., Malone, J., & Tyree, A. (1989). Prevalence and stability of physical aggression between spouses: A longitudinal analysis. *Journal of Consulting and Clinical Psychology*, 57, 263-268.
- O'Leary, K. D., & Curley, A. D. (1986). Assertion and family violence: Correlates of spouse abuse. *Journal of Marital and Family Therapy*, 12, 281-289.
- O'Leary, K. D., Vivian, D., & Malone, J. (1992). Assessment of physical aggression in marriage: The need for a multimodal method. *Behavioral Assessment*, 14, 5-14.
- Olweus, D. (1979). Stability of aggressive reaction patterns in males: A review. *Psychological Bulletin*, 85(4), 852-875.
- Pan, H. S., Neidig, P. H., & O'Leary, K. D. (1994a). Predicting mild and severe husband-to-wife physical aggression. *Journal of Consulting and Clinical Psychology*, 62, 975-981.
- Pan, H. S., Neidig, P. H., & O'Leary, K. D. (1994b). Male-female and aggressor-victim differences in the factor structure of the Modified Conflict Tactics Scale. *Journal of Interpersonal Violence*, 9, 366-382.
- Rathus, J., O'Leary, K. D., & Meyer, S. L. (in press). Attachment, proximity control, and wife abuse. *Journal of Family Violence*.
- Rosenbaum, A., Gearan, P., & Ondovic, C. (1997, August). Completion and recidivism among court-referred and self-referred batterers in a psychoeducational group treatment program. In A. Rosenbaum (Chair), *Batterers treatment: Strategies, issues, and outcomes*. Symposium presented at the annual meeting of the APA, Chicago, IL.
- Rosenbaum, A., & Maiuro, R. (1989). Eclectic approaches in working with men who batter. In P. L. Caesar & L. K. Hamberger (Eds.), *Treating men who batter: Theory, practice, and programs* (pps. 165-195). New York: Springer.
- Rosenbaum, A., & O'Leary, K. D. (1986). The treatment of marital violence. In N. S. Jacobson & A. S. Gurman (Eds.), *Clinical handbook of marital therapy* (pp. 385-407). New York: Guilford Press.
- Saunders, D. G. (1996). Feminist-cognitive-behavioral and process-psychodynamic treatments for men who batter: Interaction of abuser traits and treatment models. *Violence and Victims*, 11, 393-414.
- Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family*, 38, 15-28.
- Spitzer, R. L., Williams, J. B. W., Gibbons, M., & First, M. B. (1992). The Structured Clinical Interview for *DSM-III-R* (SCID) I: History, rationale, and description. *Archives of General Psychiatry*, 49, 624-629.
- Stamp, G. H., & Sabourin, T. C. (1995). Accounting for violence: An analysis of male spousal abuse narratives. *Journal of Applied Communication Research*, 23, 284-307.
- Stets, J. E., & Straus, M. A. (1990). Gender differences in reporting marital violence and its medical and psychological consequences. In M. A. Straus & R. J. Gelles (Eds.), *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families* (pp. 151-165). New Brunswick, NJ: Transaction Publishers.
- Straus, M. A. (1979). Measuring intrafamily conflict and violence: The Conflict Tactics (CT) Scale. *Journal of Marriage and the Family*, 41, 75-78.

- Straus, M. A., & Gelles, R. J. (Eds.). (1990). *Physical violence in American families: Risk factors and adaptations to violence in 8,145 families*. New Brunswick, NJ: Transaction Publishers.
- Straus, M. A., Gelles, R. J., & Steinmetz, S. K. (1980). *Behind closed doors: Violence in the American family*. Garden City, NY: Anchor Books/Doubleday.
- Tolman, R. (1989). The development of a measure of psychological maltreatment of women by their male partners. *Violence and Victims*, 7, 159–177.
- Tolman, R. (1999). The validation of the psychological maltreatment of women inventory. *Violence and Victims*, 14, 25–35.
- Vivian, D., & Heyman, R. E. (1996). Is there a place for conjoint treatment of couple violence? *In Session*, 2, 25–48.
- Vivian, D., & Langhinrichsen-Rohling, J. (1994). Are bi-directionally violent couples mutually victimized? A gender sensitive comparison. *Violence and Victims*, 9, 107–124.
- Vivian, D., & Malone, J. (1997). Relationship factors and depressive symptomatology associated with mild and severe husband-to-wife physical aggression. *Violence and Victims*, 12, 3–18.
- Whitehurst, G. R. (1984). Interrater agreement for journal manuscripts. *American Psychologist*, 39, 22–28.
- Yllö, K. (1993). Through a feminist lens: Gender, power, and violence. In R. J. Gelles & D. R. Loseke (Eds.), *Current controversies on family violence* (pp. 47–62). Newbury Park, CA: Sage.

RECEIVED: March 31, 1998

ACCEPTED: December 21, 1998

# **Feminist-Cognitive-Behavioral and Process-Psychodynamic Treatments for Men Who Batter: Interaction of Abuser Traits and Treatment Models**

**Daniel G. Saunders**  
*University of Michigan*

At a community-based domestic violence program, 218 men with a history of partner abuse were randomly assigned to either feminist-cognitive-behavioral or process-psychodynamic group treatments. The treatments were not hypothesized to differ in outcome. However, men with particular characteristics were expected to have lower recidivism rates depending on the type of treatment received. Treatment integrity was verified through audio-taped codings of each session. The partners of 79% of the 136 treatment completers gave reports of the men's behavior an average of 2 years post-treatment. These reports were supplemented with arrest records and self-reports. Rates of violence did not differ significantly between the two types of treatment nor did reports from the women of their fear level, general changes perceived in the men, and conflict resolution methods. However, interaction effects were found between some offender traits and the two treatments. As predicted, men with dependent personalities had better outcomes in the process-psychodynamic groups and those with antisocial traits had better outcomes in the cognitive-behavioral groups. The results suggest that more effective treatment may occur if it is tailored to specific characteristics of offenders.

Since the 1970s treatment programs for men who batter have proliferated, but tests of their effectiveness have not kept pace. Evaluations rarely include adequate comparison or control groups. For example, of the 26 studies reviewed by Hamberger and Hastings (1993) and Tolman and Edleson (1992), only 7 had comparison groups and only one of them had a true experimental design. Moreover, the posttreatment follow-up periods have generally been brief and do not always rely on reports from the men's partners, which are the most reliable sources of information. These studies and the problems with their methods have been carefully reviewed elsewhere (Hamberger & Hastings, 1993; Holtworth-Munroe, Beatty, & Anglin, 1995; Rosenfeld, 1992; Tolman & Edleson, 1995; Saunders & Azar, 1989).

The most commonly evaluated method is cognitive-behavioral, primarily in a men's group format and usually combined with gender role resocialization and methods to reduce male dominance. Only one study could be founded which experimentally

compared this approach with another approach. Edleson and Syers (1990,1991) compared: (1) a minimally structured self-help model with a peer facilitator and professional consultant, (2) a structured educational model (cognitive-behavioral) with regular readings and assignments, and (3) a combined model that presented material in less detail than the educational model and allowed more time for work on individual problems. Reports from 52% of the partners, or 8% of the men 6 months after treatment, nonsignificantly favored the education and combined groups. At 18 months posttreatment, reports from 46% of the women or the men showed even less difference among treatment approaches. Other studies used quasi-experimental designs or had small samples (e.g. Dutton, 1986; Harrell, 1990; Stosny).

In the study reported here, I attempted to improve on previous evaluations by obtaining a higher rate of response during follow-up and by ensuring that the treatments were applied according to their stated goals. The study also tested more theoretically distinct treatment models than the Edleson and Syers comparison. Risk factor research on domestic violence, which supports a number of theoretical explanations for domestic violence (Van Hasselt, Morrison, Bellak, & Hersen, 1988) suggested two different approaches for comparison. This research consistently shows that boys who witness violence between their parents or who are abused themselves are more likely to be spouse abusers when they grow up (Hotaling & Sugarman, 1986). Psychological abuse of boys by their parents also seems to be a risk factor (Dutton, van Ginkel, & Starzomski, 1995). However, different theories can be used to explain the intergenerational transmission of violence. Recent reviews of the research include genetic, social learning, and attachment theories (Holtzworth-Munroe & Stuart, 1994).

Principles of social learning theory, such as imitation and modeling, are frequently offered to explain the transmission (O'Leary, 1988). In addition, the lack of adequate role models means that interpersonal skills will be lacking in these men and there is evidence that they have such skill deficits (reviewed in Holtzworth-Munroe, 1992; Tolman & Bennett, 1990). They may also learn cognitive distortions from their fathers or have trauma-induced beliefs about themselves that lead to cognitive distortions. Thus cognitive therapies are often used with these men. These individual level explanations of learning are often combined with sociocultural explanations. Some feminist theories, for example, maintain that these individual factors will most likely lead to violence in cultures that support violence against women. There is evidence from cross-cultural studies that patriarchal norms and structures are risk factors for abuse (Levenson, 1989; Yllo, 1984). Many treatment programs integrate behavioral, cognitive, and feminist theories in their treatment approaches (e.g., Ganley, 1989) and one of the models tested here was such a combined Feminist-Cognitive-Behavioral Treatment (FEBT) approach. It was developed over a number of years and is typical of many programs throughout the country (e.g., Edleson, 1984; Saunders, 1984). It combined skills-training and gender role resocialization in a highly structured format. Brief lectures, demonstration role-plays, behavioral rehearsal, and homework assignments were the primary methods. More details about treatment methods are presented in the Method section. This model assumed that the violence is caused by: (1) behavioral skills deficits that reduce men's ability to state needs and feelings directly, thus repressing anger or leading to its immediate expression; (2) cognitive skills deficits that generate anger and/or rationalizations internally through cognitive distortions and negative self-talk; and (3) cultural norms and structures that support male dominance and fail to punish woman abuse.

The risk factor research can also support a psychodynamic perspective. The childhood abuse commonly witnessed or experienced by these men can have several outcomes with psychodynamic explanations. The explanations center on attempts to resolve childhood trau-

mas and subsequent development of various attachment disorders. This theoretical model assumes that the men will identify with the aggressor in order to feel more powerful, perpetrate the same trauma onto others in order to gain a sense of control over the trauma, and/or defend against painful memories through anger and aggression. There is growing evidence that many of these men develop PTSD symptoms related to childhood trauma (Dutton, 1995) and a variety of personality disorders, including borderline/schizoid, antisocial/narcissistic, and passive-dependent/compulsive disorders (Hamberger & Hastings, 1986). Psychodynamic theories can also be placed in a cultural context, in particular the impact of patriarchy on male socialization and the channeling of vulnerable feelings into anger (Scher & Stevens, 1987).

The psychodynamic interpretation of the risk factor research led to the development of a Process-Psychodynamic Treatment (PPT) model. It was much less structured than the FCBT model because it assumed that the most effective learning occurs through the process of supportive, nondidactic group relationships (Jennings, 1987). The model assumed that violence was caused primarily by childhood traumas experienced by the men (e.g., witnessing abuse of a loved one and/or being abused) and that the latter displace anger about these traumas is displaced onto adult relationships. The leaders of this approach try to create a supportive environment in which the men can reexperience childhood traumas, grieve their losses, give up control over others, and learn to empathize with others (Whitfield, 1987). "Insight" models of this type have been criticized for failing to confront the violence directly and overemphasizing support and empathy, which may reinforce abusers' rationalizations; insight approaches may also take too long to work (Adams, 1988). Such reservations about unstructured, insight models led to the development of feminist-cognitive-behavioral models (Saunders, 1984), yet we believed that all models needed to be empirically tested.

One purpose of this study was to test the relative effectiveness of these two models for preventing the recurrence of men's violence against their partners. Although cognitive-behavioral methods had been evaluated previously, process-psychodynamic methods had not been. Neither treatment was predicted to be superior on the main outcome measure: the partners' reports of the abusers' behavior after treatment. Comparisons of cognitive-behavioral and insight approaches often find no differences on major outcome variables. For example, Deffenbacher and his colleagues (Deffenbacher, McNamara, Stark, & Sabadell, 1990) found no differences between cognitive-relaxation treatment and a process-oriented approach to anger management in a well-controlled study with college students.

Another purpose of the study was to test the proposition that offenders with particular traits would have better outcomes depending on the type of treatment they received. Studies of rehabilitation programs for offenders of all types show poor results unless they are matched with appropriate treatments. One meta-analysis (Andrews et al., 1990) showed that positive outcomes only occur when the styles and modes of treatment are matched with client needs and styles of learning.

Competing theories about domestic violence may be reconciled at least in part through evidence of different types of abusers. Abusers can be differentiated along a number of variables, including the extent of childhood victimization, type of personality disorder, and attitudes about women. Three types generally emerge (Holtzworth-Munroe & Stuart, 1994): (1) Men who experienced the most severe childhood physical abuse tend to develop antisocial traits including generalized aggression, substance abuse, and proviolence norms. They show little or no remorse, have the most rigid views of gender roles, and can be theorized to have a dismissing attachment style. (2) Men who experienced the most severe

parental rejection appear to develop borderline traits (Dutton & Starzomski, 1993) and tend to be the most emotionally abusive, are emotionally volatile, and have the highest dependency needs. They probably have preoccupied or ambivalent attachment styles. (3) Men who experienced the least amount of childhood trauma tend to be compulsive and lacking in communication skills. They have low to moderate dependency needs and may have secure or preoccupied attachment styles.

The feminist-cognitive-behavioral model seems best suited to men with antisocial traits. Although they were most severely abused as children, they are not likely to be aware of the need to resolve their traumas and are unlikely to respond well to relationship-based, insight approaches. They probably need the most work on skills-training and attitude change. The process-psychodynamic model seems best suited for men with moderate to high levels of dependency needs because they are much more likely to engage in group process and methods for enhancing self-awareness. Therefore, in this study, it was predicted that the more dependent the personality, the more likely there would be better outcomes in the process psychodynamic group. Conversely, the more antisocial the personality, the more likely there would be better outcomes in the feminist-cognitive-behavioral group. Outcome was based on the partners' reports of abuse, conflict-resolution method, fear, and general changes in the men many months after they ended treatment.

## METHOD

### Sample

**Men.** Those who agreed to participate in the experiment ( $n=218$ ) were recruited from men who had been assessed and accepted for treatment at a family counseling agency that was certified as an outpatient mental health clinic. Most of the men had been referred by a deferred prosecution program (17%) or probation department following prosecution (59%). The others were referred by social service agencies, attorneys, friends, family members, or themselves. The average age of the men was 32.4 ( $SD=8.3$ ). Fourteen percent were African-American, 3% were Hispanic, 4% Native American, 1% Asian, and the remainder were Euro-American (78%). Their average income was \$13,435 per year ( $SD=\$10,162$ ). All but 18% of the men had graduated from high school, with 23% having some college, 11% having college degrees, and 2% having attended graduate school. Their average years of education was 12.6 ( $SD=1.9$ ). Men who completed treatment ( $n=136$ ), who are the focus of this report, were more likely to be Euro-American (84%), to have higher incomes (\$14,540), and to have more years of education (13.0). During an intake interview, 53% of the men reported being punished by a parent more severely than the use of a slap or spanking. Almost half (43%) reported verbal abuse from a parent.

**Women.** Attempts were made to contact all of the partners of the men who completed the assessment phase and agreed to participate in the experiment (procedures described below). Of the 218 men assigned to a treatment condition, 199 of the partners could be located and they were contacted. There were several purposes for these contacts: (1) to inquire about their willingness to participate in pretreatment and posttreatment interviews; (2) to inform them of the emergency, counseling and legal services available to them in the community; (3) to explain what might occur while their partners were in treatment and of the need not to view their partners' treatment as a panacea; and (4) to initiate a safety plan if needed. Pretreatment interviews were designed to last 30 to 45 minutes and posttreat-

ment interviews were designed to last 10 to 20 minutes. The women were given a choice of telephone or in-person interviews or of answering a mailed questionnaire. Procedures similar to those used by other researchers (Parker & Ulrich, 1990) were developed and implemented to protect the safety of the women and the interviewer regardless of the data collection method. The original design called for data collection at the assessment phase, the beginning of treatment, half-way through treatment, and 3, 6, 9, 12, and 18 months after treatment. The focus was shifted to 18-month follow-ups and beyond because of initial difficulties in locating women and because of evidence that many men are likely to be violent beyond the first 6-month follow-up phase (Dunford, 1992).

The interviews covered psychological and physical abuse, level of fear, general changes in the men, and conflict resolution methods (described later). The women were asked about any incidents of abuse since the end of treatment or since their last posttreatment interview. In the initial interview we also obtained the names and telephone numbers of friends and relatives likely to know their whereabouts during the follow-up phase. The women were paid between \$10 and \$30 for each interview, depending on its length and time since treatment. These and other methods of locating battered women and providing incentives for participation had been used successfully by other researchers (Rumptz, Sullivan, Davidson, & Basta, 1991). To increase response rates during long-term follow-up of 18 months and beyond, a short version of the questionnaire was offered (described below).

Of the men who completed treatment ( $n=136$ ), 79% of their partners were located and agreed to participate during the follow-up period (86% ( $n=55$ ) in the FCBT condition and 72% ( $n=52$ ) in the PPT condition). About half of the women provided information once (46%), but a large proportion (42%) provided information 4 to 6 times (41%, FCBT; 45% PPT). All but 5% of the final data gathering occurred 18 to 54 months after treatment (96% FCBT; 94% PPT). The five cases with 3-12-month interviews all reported violence from at least one source and were not prioritized for further follow-up. Most interviews occurred between 18-23 months posttreatment (56%). The distribution of interviews in 6-month segments over the 54-month period was nearly identical between the two conditions. The average length of follow-up for FCBT was 26.0 months ( $SD=11.2$ ) and for PPT was 24.6 months ( $SD=9.4$ ), a nonsignificant difference ( $t=.70$ ,  $p=.48$ ). For those with interviews at 2 or more years posttreatment, the average follow-up periods again did not differ between the conditions (FCBT:  $M=36.0$  months,  $SD=10.0$ ; PPT:  $M=34.7$  months,  $SD=7.8$ ;  $t=.45$ ,  $p=.65$ ).

## Procedures

**Intake Procedures.** Although most of the men were referred by the criminal justice system, all men were required to call the program for an appointment. They normally attended four to six individual intake sessions. In addition to obtaining a comprehensive history of past help seeking, substance abuse, suicide potential, childhood violence, relationship violence and other areas, these sessions included attempts to increase the man's motivation for change, decrease his minimizing about abusive behavior, and develop a control plan. When appropriate, the partner was invited to rehearse a "time-out" procedure (Sonkin, 1989) at the end of the assessment phase. Only a small percentage of the men were screened out because their violence was directed only at nonintimates or because the intake worker decided they could not benefit from group treatment due to severe psychopathology, developmental disabilities, or complete denial of problems. The men also completed a series



of self-administered questionnaires on attitudes, behavior, and affect (described below). Because the groups were close-ended, some of the men had to wait 2 to 8 weeks to join a treatment group. While waiting, they attended an orientation group every other week for 2 hours each.

**Random Assignment.** Once a man was determined eligible for treatment, the intake worker explained the experiment to him and he was given an informed consent form to read. Clients were given the option of participating in the experiment or completing a regular course of treatment. The regular treatment consisted of 12 sessions of feminist-cognitive-behavioral group treatment followed by 20 sessions of a mutual support group. Twenty-two men chose the regular condition. The remaining 213 men who were eligible and agreed to participate in the experiment were randomly assigned to one of the two treatments. Five other men were included in the analysis who were assigned to groups based on other criteria. For example, one of the men placed in the PPT groups already had assertiveness training; two other men, both voluntary referrals, had a strong preference for one type of treatment. Supplementary data analyses excluded these five men.

Although the men were to be assigned simultaneously to each condition from the list of eligible men, two of the nine assignments were not simultaneous. One of the PPT groups had too few members to be viable. The group was stopped and then started 3 months later with additional members who were also randomly assigned to that condition. One FCBT group had to be composed only of men with health insurance because the only leader available to start a group was an outside contractor the agency could not otherwise afford. The noninsured men on this waiting list were placed in an FCBT group a month later. Thus the men were still assigned to the proper condition. These problems with randomization should not pose a major threat to validity (e.g., history) because multiple groups of both conditions occurred over a 3-year period.

Of the 218 men assigned to groups, 178 (82%) attended the first group session, 91 in the FCBT condition and 87 in the PPT condition.

**Treatments.** Both types of treatments used close-ended groups of 20 weekly sessions lasting 2.5 hours each. The FCBT condition followed a highly structured format. Agendas and homework assignments were included in each session (agendas are available on request from the author). Each session included a didactic section on communication and cognitive skills, relaxation/desensitization training, consciousness raising about sex roles and violence against women, and behavioral or cognitive rehearsal. It included the major ingredients of other FCBT models (e.g., Edleson & Tolman, 1992; Stordeur & Stilles, 1989).

The PPT did not use agendas but instead followed several phases over the 20 weeks: building trust and a sense of safety; uncovering the childhood traumas and reconnecting with traumatic childhood events; mutual support and awareness of hurt and fear; building awareness of alienation from self and others; transferring lessons about reactions to abuse to current relationships and dealing with termination. Although the group was much less structured than the FCBT condition, it was more structured than most psychotherapy groups. This model is described in detail elsewhere (Browne, Saunders, & Staecher, in press). Several handouts were used from Gil's (1983) and Whitfield's (1987) work on recovery from child-abuse and neglect.

**Treatment Integrity.** A potential problem with an experiment's internal validity is the misapplication of methods by the leaders. Audiotape recordings were made of each group session to aid in the assessment of treatment validity. These tapes were also used by group supervisors for supervisory purposes. The author created 58 codes in four areas: leader methods, group content, time focus of discussions (e.g., childhood, past month, here and

now), and relationship focus of discussions (e.g., parents, partner, other group members). The final list of code categories ( $n=25$ ) was based on ratings by the two primary supervisors regarding the relative importance and predicted time that would be spent on their respective approaches. Graduate students with no investment in the outcome of the treatments were trained to code the tapes. Four 8-minute audio segments were coded from the beginning, middle, and end of each session. A total of 76 hours from each condition was coded, or 17% of total group time. Each audio segment was reviewed twice for the occurrence of each code. No attempt was made to measure the duration of each behavior but only whether it occurred or not. Three-and-a-half-hours of the tapes (2.1% total tape time) were coded by two raters. Interrater reliability using percent agreement was adequate (70%-81%), except for two areas, leader lecturing and self-disclosure, which were below 40%.

The frequencies of each category by treatment condition are shown in Table 2. Relaxation training and work on coping thoughts occurred almost exclusively in FCBT, as expected. Building emotional awareness, "becoming aware of feelings", on the other hand, occurred at about equal rates. The focus on emotional safety occurred much more often in PPT than in FCBT and the focus on childhood loss and abuse occurred almost solely in the PPT groups. Behaviors characteristic of the FCBT condition, leader and member role-playing, did not occur in the PPT condition. Advise-giving from both leaders and members showed some overlap as expected, yet occurred more than twice as often in the FCBT groups. The same was true for leader lecturing. Self-disclosure also overlapped, yet occurred twice as often in PPT groups, as expected. The PPT focused three times as often on the men's parents and nearly three times as much on group members or leaders. Thus, the treatment delivery appeared to be consistent with the two theoretical frameworks. These findings are consistent with a study of cognitive-behavioral and psychodynamic processes (Jones & Pulos, 1993) in which the cognitive-behavioral techniques emphasized didactic methods, discussion of cognitions, and explicit advice; psychodynamic techniques emphasized memories or reconstructions of childhood, linking feelings to past situations and discussion of the therapy relationship.

**Treatment Completion.** Completion for this study was defined, as it was by program policy, as attendance at 16 out of 20 sessions. Leader judgments about success, regardless of treatment length, were also taken into account and used in supplementary analyses.

**Group Leaders.** All of the primary group leaders had extensive experience conducting treatment groups of male offenders. They were clinical social workers or psychologists. All but one group had a coleader. Coleaders were either other social workers or psychologists,

TABLE 1. Source of Recidivism Reports

|                                | FCBT | PPT |
|--------------------------------|------|-----|
| Woman, man, and arrest records | 42%  | 38% |
| Woman and arrest               | 29%  | 25% |
| Woman and man                  | 5%   | 6%  |
| Man and arrest                 | 11%  | 11% |
| Woman only                     | 8%   | 4%  |
| Arrest only                    | 5%   | 14% |
| Man only                       | 0%   | 1%  |
| No report                      | 0%   | 1%  |

$$\chi^2 = 6.1, p = .53.$$

or master's degree social work interns who had observed at least one series of group sessions. The first two of the PPT groups also had peer group coleaders. We discontinued their use because they had completed the FCBT program and that was their primary orientation for helping the men. An analysis of session tapes with and without these leaders revealed no significant differences in leader methods or the focus of group discussions. Client ratings of helpfulness and support from postgroup questionnaires seemed to favor the professionally led groups, but statistical tests could not be conducted due to small samples.

Of the nine FCBT groups there were seven male-female cotherapy teams, one male-male team, and one male-led group. Of the nine PPT groups, there were seven male-female teams and two male-male teams. Each treatment condition had an African American male leader and the rest of the leaders were Euro-American. None of the leaders in either condition crossed over to lead the other condition. However, one worker originally trained to conduct FCBT groups became a PPT group leader at the start of the experiment. Leaders were chosen partly for their preference for each theoretical orientation.

The supervisor of the FCBT groups had a master's degree in social work and 7 years experience working with men who batter. Three clinicians helped supervise the PPT groups. Each one had more than 5 years experience treating male offenders (assaulters or sex offenders). One was a clinical psychologist and two were clinical social workers. None of them had supervised groups for men who batter because this treatment model had never been used before in this community. The primary supervisor and developer of the model had extensive individual and group experience treating male sex offenders.

## **Community Context**

The groups met at an established domestic violence program within a nonprofit family service agency. The program participated in a county-wide plan to coordinate victim and offender services with the response of the criminal justice system. The major law enforcement jurisdictions had pro-arrest policies and the prosecutor had a first offenders' program and a victim support program. Probation officers and first offender program staff were trained in the field of domestic violence. As described earlier, 59% of the men were referred by the courts, 17% more were under deferred prosecution, and most of the remainder were from social service agencies.

## **Design**

The study was conducted as a randomized field experiment. As with many such experiments, conditions often become nonequivalent due to attrition during and after the treatment phase (Cook & Campbell, 1979). Therefore, checks were made on the equivalency of the treatment completers from the two conditions on many behavioral, attitude, and demographic variables.

## **Equivalency of Conditions**

Despite treatment attrition rates of 38% for the FCBT condition and 24% for PPT, the random assignment was apparently not compromised. There were no significant differences between the groups on pretreatment measures of personality, attitudes, depression, anger, partner reports of violence, or number of arrests described below. There were also no differences in age, years of education, income, years in the relationship, use of prior treatment, or mandatory referral. The FCBT condition had more Euro-American clients (91%

vs. 77%)(chi-square = 8.3;  $p = .04$ ), and tended to have more divorced (44% vs. 29%) but fewer separated (7% vs. 13%) clients.

Despite the near equivalence of the groups, there were differential predictors of attrition. Younger, less educated men, who had not been victims of child abuse tended to drop out of the FCBT condition. Voluntarily referred men who had witnessed parental violence were more likely to drop from the PPT condition (Chang & Saunders, 1993).

## Recidivism Measures

Recidivism was primarily measured by the women's reports and was supplemented by men's reports and arrest records.

**Women's and Men's Reports of Violence.** An expanded version of the Conflict Tactics Scale (CTS) (Straus & Gelles, 1990) was used for pre-, post- and follow-up reports from the men and women. Items were added on nonviolent threats, using a car recklessly, and sexual abuse (Saunders, 1992). The version contained 12 psychological abuse items and 14 physical abuse items. A short version, offered to women who did not want to complete the long version, condensed the items on abuse to three items and asked for absolute frequencies. The items were: "(a) verbally or emotionally attack you, including insulting, swearing, threatening to leave you, saying you couldn't see certain people; sulking, damaging property, or similar behavior; (b) physical force against you, including threatening to hit you or throw something at you; pushing, carrying, restraining, or grabbing; slapping you; driving recklessly to frighten you; throwing an object at you; kicking you or hitting you with a fist; throwing you bodily; physically forcing sex on you; hitting or trying to hit you with something; (c) beat you up (multiple blows), choke you, make threats with a weapon, or used a weapon against you." Marital status, periods of separation, and how disagreements were handled were also included on the questionnaire. Inquiries about the desire for service for the woman or her children were also made. The focus of this report is on the occurrence of any of the physically abusive behaviors after treatment. The primary source of information was the partner reports 18 to 54 months after treatment. Reports from the men and from official arrest records were used when the women's reports were not available or if the men's reports or the arrest records revealed any violence when the women reported none.

The men were mailed questionnaires containing the expanded CTS 12 months after treatment. Just over half of them returned the questionnaire (FCBT: 58%; PPT: 54%). The report of only one man was used exclusively (no partner or arrest report). He was in a PPT group and reported no violence. In two cases the men's reports proved very useful in detecting recidivism because they reported violence and there were no partner reports and the arrest record showed no arrests.

**Arrest Records.** Pretreatment and posttreatment arrest records of most of the men were obtained from the state's criminal justice computer (FCBT: 87%; PPT: 89%). The post-treatment time period ranged from 2.0 to 4.6 years. Arrests were placed in categories: property, financial, person, criminal justice system, weapons, traffic, nontraffic substance abuse, and "other." The focus in this study is on crimes against persons. Most of these were listed as "domestic battery," others were simply labeled "battery." Early in the experiment the state law did not have a specific domestic battery statute, but in some cases it seemed likely the "battery" was often directed at a partner because of the combination of offenses, for example, "concealed weapon, battery, disorderly conduct, attempted sexual assault" in one case, and "false imprisonment, restraining order violation, endangering safety with

weapon" in another. A case of "resisting and obstructing an officer" was also counted as recidivism.

The arrest records were useful because in 19 cases of arrests against persons the woman's report was not available, she had no contact with her partner, or she reported no violence.

**TABLE 2. Treatment Integrity: Frequency of Behaviors Over 76 Hours of Audio Recording**

|                                                                                             | FCBT<br>Feminist-Cognitive-<br>Behavioral | PPT<br>Process-<br>Psychodynamic |
|---------------------------------------------------------------------------------------------|-------------------------------------------|----------------------------------|
| <b>Content:</b>                                                                             |                                           |                                  |
| Progressive relaxation                                                                      | 54                                        | 8                                |
| Using coping thoughts (pos. beliefs about one-self—attributes, abilities, behavior, etc.)   | 63                                        | 3                                |
| Definitions of verbal and physical abuse (including marital rape)                           | 16                                        | 15                               |
| Becoming aware of feelings                                                                  | 39                                        | 36                               |
| Childhood losses and rejections (includes psychological abuse and being child of alcoholic) | 0                                         | 62                               |
| Childhood experience with violence (seeing abuse or being abused)                           | 5                                         | 38                               |
| Emotional safety in group                                                                   | 7                                         | 50                               |
| $\chi^2 = 200.3, p < .0001.$                                                                |                                           |                                  |
| <b>Methods:</b>                                                                             |                                           |                                  |
| Leader role-play (modeling: at least 2 people in verbal exchange)                           | 26                                        | 0                                |
| Member role-play (rehearsal)                                                                | 83                                        | 1                                |
| Advise (giving suggestions & directions by member or leader)                                | 240                                       | 128                              |
| Lecture (provide information about skills, concepts & problems)                             | 172                                       | 45                               |
| Self-disclosure (by leader or member)                                                       | 95                                        | 215                              |
| $\chi^2 = 220.91, p < .0001.$                                                               |                                           |                                  |
| <b>Time focus of discussions:</b>                                                           |                                           |                                  |
| Distant past (0 to 18 years)                                                                | 19                                        | 101                              |
| Near past (18 yrs to 1 mo. ago)                                                             | 170                                       | 283                              |
| Most recent (past month)                                                                    | 405                                       | 411                              |
| Here & now (events in group)                                                                | 145                                       | 164                              |
| Near future (next month)                                                                    | 91                                        | 105                              |
| Distant future (over 1 month)                                                               | 59                                        | 64                               |
| $\chi^2 = 59.1, p < .0001.$                                                                 |                                           |                                  |
| <b>Relationship focus of discussions</b>                                                    |                                           |                                  |
| Intimate partner                                                                            | 299                                       | 307                              |
| Man's parents                                                                               | 34                                        | 101                              |
| Co-worker                                                                                   | 41                                        | 21                               |
| Group member or leader                                                                      | 69                                        | 198                              |
| Self                                                                                        | 351                                       | 443                              |
| $\chi^2 = 73.5, p < .0001.$                                                                 |                                           |                                  |

In some of these cases, he was probably violent toward a new partner. Although the men's report and the arrest report were congruent in 10 cases, in two cases of arrest the men reported no violence. A limitation of this measure is that in nine cases in which no arrest was reported, we could not verify if either partner was living in the state.

The response rate of the women during the follow-up period was 79%. If one adds the remaining cases in which either an arrest occurred ( $n = 8$ ) or the man reported violence ( $n = 1$ ), the "conservative" response rate is 85%. If one includes the presence of all three reports (given the problems noted above), the "liberal" response rate is 99%. The source of report (woman, man or arrest) or the various combinations did not differ significantly between conditions. This information appears in Table 1.

## Women's Measures

**Fear.** The women were asked the extent to which they feared physical abuse from their partners before treatment and at each follow-up point: "In general, I fear physical abuse from my partner: not at all, a little bit, a moderate amount, a great deal." A second item substituted "psychological abuse" for physical abuse.

**Conflict Resolution.** A single item was used to assess general conflict style before and after treatment: "When disagreements arise, do they generally result in: man giving in, woman giving in, neither giving in, agreement by mutual give and take."

**General Changes.** The women were asked two open-ended questions during the follow-up: (1) "During or since the group, have you noticed any positive changes in your partner?" ; (2) During or since the group, have you noticed any negative changes in your partner." In most cases more than one change was noted. Cases were classified as: positive only, negative only, or a mixture of the two. The frequencies of positive and negative changes for each case were also recorded.

## Men's Measures.

A packet of self-report questionnaires were administered to the men during the intake phase and again between the last group session and an exit interview with a counselor. This report will focus on the use of these measures in determining the success of randomization and trait-treatment interactions.

**Millon Clinical Multiaxial Inventory (MCMI-1).** The MCMI was used to assess the personality traits and disorders of the men (Millon, 1983). It contains 175 items that measure eight personality and character disorders (Axis I), three chronic and dysfunctional personality disorders (Axis II) and nine circumscribed or transient clinical syndromes. The MCMI shows good internal-structural validation and external validity with many other measures. It contains corrections for psychological defensiveness, self-deprecation, and denial tendencies. The dependent and antisocial-aggressive scales were the most relevant in this study. All 20 scales were factor-analyzed to reduce the data. The five resulting factors were used to further test the hypotheses. They were: (1) dependent/somatoform; (2) drug-abuse/alcohol abuse/ narcissistic/ hypomanic/ antisocial; (3) avoidant/borderline/anxious/depressed; (4) paranoid; and (5) compulsive/passive-aggressive. The traits with the highest loadings are placed first in the above lists.

**Relationship Satisfaction.** A short, 11-item version of a marital satisfaction scale was used and relabelled the "Relationship Inventory" (Roach, Frazier, & Bowden, 1981). The highest loading items were chosen from the original scale. The original scale shows very

high concurrent validity and internal reliability, and nonsignificant correlations with social desirability.

**Beliefs About Woman Abuse.** The Inventory of Beliefs about Wife Beating (Saunders, Lynch, Grayson, & Linz, 1987) was used to measure the men's beliefs and attitudes. It contains 31 items in five subscales: wife beating is justified, wives gain from the abuse, help should be given, the offender is responsible, and the offender should be punished. Evidence exists for its concurrent and known groups' validity using various populations. The first three subscales can be combined into a scale of sympathy toward battered women that has very good internal reliability.

**Self-Esteem.** The Rosenberg Self-Esteem Scale was used. This is a 10-item measure of self-esteem that shows good construct and concurrent validity and high internal reliability (Fleming & Courtney, 1984). The version used here had a 4-point response format from "strongly disagree" to "strongly agree."

**General Hostility.** The Buss-Durkee Hostility Inventory consists of 66 items in seven subscales (Buss & Durkee, 1957): assault, indirect hostility, irritability, negativism, resentment, suspicion, and verbal hostility. It appears to have a two-factor structure, one emphasizing resentment and suspicion and the other aggressive behaviors. Almost all of the subscales appear to discriminate well between violent and nonviolent populations. The internal reliability of some of the subscales is not very high.

**Traditional Views of Women's Roles.** The 15-item version of the Attitudes Toward Women Scale was used (Spence & Helmreich, 1978). The response format is from "strongly agree" to "strongly disagree." Its internal reliability with college men is .89 (alpha). It demonstrates construct validity through its ability to differentiate males and females and older and younger persons; its ability to predict profeminist reactions to competent women; and its correlation with acceptance of gender stereotypes (Spence & Helmreich, 1978).

**Democratic Decision Making.** A short version of the Power Decision Index (Blood & Wolfe, 1960) was used to measure the extent to which the man or the woman has the final say in five areas of marital decision making. The greatest weight is given if the couple shares decision making and the least is given if either one is dominant. A modified version of the scale, used in the first national study of family violence (Straus, Gelles, & Steinmetz, 1980) was used here.

**Level of Conflict.** This construct was measured with the Marital Conflict Index. It was used in the first national study of family violence (Straus, Gelles, & Steinmetz, 1980). Respondents are asked to rate five areas in the relationship (managing money, affection and sexual relations, household chores, social activities, and children) on the frequency of agreement in the past year, from "always" to "never."

**Anger Toward Partner.** A spouse-specific version of the Novaco Anger Index (Novaco, 1975) was used. It uses a 5-point scale to indicate reactions to situations, from "very little anger" to "very much anger." The index showed positive changes following cognitive and relaxation treatments. Twenty of the interpersonal items from the original 80-item scale were changed to "partner." In previous studies with men who batter the internal reliability coefficient was .89 (Saunders & Hanusa, 1986).

**Jealousy.** A measure of romantic jealousy developed by White (1977) was used. It contains 6 items. It has high internal reliability and correlates as expected with dependency on the relationship.

**Depression.** The Beck Depression Inventory was used (Beck, 1961). It is a 21-item measure covering somatic complaints, guilt, pessimism and indecisiveness. The split half reli-

ability is .86. Criterion validity has been demonstrated based on the inventory's correlations with ratings of patients.

**Adjustments for Social Desirability.** The attitude and affect measures were adjusted for the tendency of the men to answer in a socially desirable manner. A 10-item version of the Marlowe-Crowne Scale with a 7-point Likert scale (Greenwald & Satow, 1970) was used. It has an internal reliability coefficient equivalent to the original scale.

## RESULTS

### Expectation Effects

Because the experiment was conducted within a program that had used only one of the methods for about 10 years, the intake workers or others in the agency could have conveyed positive expectancies about the FCBT group that might lead to placebo effects. Analysis of evaluations completed by the men after the first four group sessions showed that positive expectations of change were high for both types of treatment. The men responded to the item "As a result of this program, I expect to improve my ability to prevent my psychological and physical abuse of others." Both groups averaged between moderate and high expectations with the FCBT group somewhat higher on a five point scale ( $M = 4.5$  vs.  $4.3$ ,  $t = 2.94$ ,  $p = .003$ ). Perceived helpfulness of each session did not differ between the conditions on a 6-point scale, ranging from "not helpful at all" to "very helpful."

### Main Effects

Table 3 shows the recidivism rates for the two treatment conditions. Those who completed 16 or more of the 20 sessions are included, plus two other men whom the leaders considered successful. The first comparison is based only on the women's reports, regardless of whether they had contact with their partners. The next comparison is more stringent since it excludes women with no partner contact. Finally, an even more stringent comparison is shown with the addition of the men's reports and the arrest records. Regardless of the comparison made, the recidivism rates are almost identical between the conditions and did not differ significantly. The results were similar under more stringent criteria: the removal of cases not randomly assigned ( $n = 5$ ), the removal of cases if they completed 16 sessions instead of 17 ( $n = 9$ ), and the removal of cases the leaders judged unsuccessful ( $n = 8$ ).

**TABLE 3. Recidivism Rates for Physical Abuse After Treatment**

|                                                                                     | Feminist-Cognitive<br>-Behavioral (FCBT) | Process-Psychodynamic<br>(PPT) | $\chi^2$ | $p$ |
|-------------------------------------------------------------------------------------|------------------------------------------|--------------------------------|----------|-----|
| Women's reports (3-54<br>mos. after treatment)                                      | 30.9% (17/55)                            | 28.8% (15/52)                  | .001     | .98 |
| Women's reports only<br>if contact with partner                                     | 34.0% (17/50)                            | 33.3% (15/45)                  | .024     | .87 |
| Women's reports<br>(partner contact and/or<br>arrest record and/or<br>man's report) | 45.9% (28/61)                            | 48.5% (33/68)                  | .001     | .98 |
| Arrests: Any crime                                                                  | 26.8% (15/56)                            | 28.1% (18/64)                  | .001     | .94 |
| Arrests: Crimes against persons                                                     | 23.2% (13/56)                            | 20.3% (13/64)                  | .008     | .93 |



**TABLE 4. Nonviolent Outcomes Reported by the Women  
(At Last Posttreatment Interview)**

|                                               | Cognitive-Behavioral<br>(FCBT) | Process-<br>Psychodynamic (PPT) |               |
|-----------------------------------------------|--------------------------------|---------------------------------|---------------|
|                                               | <i>n</i> = 45                  | 41                              | $\chi^2$      |
| General changes during<br>and after treatment |                                |                                 |               |
| positive & negative                           | 56%                            | 59%                             | 0.8 <i>ns</i> |
| positive only                                 | 31%                            | 34%                             |               |
| negative only                                 | 13%                            | 7%                              |               |
|                                               | <i>n</i> = 44                  | 44                              |               |
| Fear of physical abuse                        |                                |                                 |               |
| not at all                                    | 54%                            | 52%                             | 5.5 <i>ns</i> |
| a little bit                                  | 39%                            | 32%                             |               |
| a moderate amount                             | 7%                             | 5%                              |               |
| a great deal                                  | 0%                             | 11%                             |               |
|                                               | <i>n</i> = 45                  | 40                              |               |
| Fear of psychological abuse                   |                                |                                 |               |
| not at all                                    | 42%                            | 50%                             | 3.3 <i>ns</i> |
| a little bit                                  | 36%                            | 28%                             |               |
| a moderate amount                             | 15%                            | 7%                              |               |
| a great deal                                  | 7%                             | 15%                             |               |
|                                               | <i>n</i> = 41                  | 38                              | 3.4 <i>ns</i> |
| Result of disagreements                       |                                |                                 |               |
| mutual agreement                              | 46%                            | 47%                             |               |
| neither give in                               | 32%                            | 32%                             |               |
| woman give in                                 | 15%                            | 21%                             |               |
| man give in                                   | 7%                             | 0%                              |               |

Table 3 also shows the rates of arrest for all crimes and crimes against persons. Again there were no significant differences. The average number of arrests in these categories of crime also did not differ significantly between the conditions (All crimes: FCBT  $M = .50$ ,  $SD = 1.4$ , PPT  $M = .72$ ,  $SD = 1.4$ ,  $t = -.85$ ; Crimes against persons: FCBT  $M = .09$ ,  $SD = .4$ , PPT  $M = .05$ ,  $SD = .2$ ,  $t = .67$ ). Psychological abuse rates did not differ between the two groups.

Table 4 shows other outcomes reported by the women. In response to the open-ended questions about changes in the men during and after treatment, just over half of the women reported that they observed both positive and negative changes. About a third observed only positive changes. Thirteen percent of the FCBT men's partners and 7% of the PPT men's partners reported only negative changes. These differences were not significant. The average number of positive and negative changes per case also did not differ significantly between the conditions (positive: FCBT  $M = 1.7$ ,  $SD = .8$ ; PPT  $M = 1.8$ ,  $SD = .9$ ,  $t = 1.36$ ; negative: FCBT  $M = 1.7$ ,  $SD = .8$ ; PPT  $M = 1.5$ ,  $SD = .8$ ,  $t = 1.12$ ).

The average level of fear for both groups was between "a little bit" and "a moderate amount" prior to treatment and fell to "a little bit" by the last point in the follow-up. After treatment, almost half of the women, regardless their partners' treatment, reported that disagreements were solved by "mutual give and take." Many reported that neither gave in (32% in both conditions). The differences between conditions were not significant.

Note that for all of these comparisons, information was available from about 50%-70% of the women, primarily because most of these measures were not available with the short questionnaire.

### Interaction Effects.

The interaction analysis sought to determine if men with certain traits had lower recidivism rates depending on which treatment they received. The Dependent and Antisocial-Aggressive subscales of the MCMI were used to test the interaction hypotheses in three forms: as continuous variables, as diagnoses (BR 75), and as part of factor scores with other MCMI traits. Forty percent of the men showed the presence of an antisocial personality since they had base rate scores of 75 on this subscale; 33% of the men had BR scores above 75 on the dependent subscale. These rates may be inflated because there is evidence that the MCMI-1 tends to overdiagnose. The dependent variable was recidivism of any form of physical abuse among treatment completers (16 or more sessions) based on reports from the woman, and/or her partner, and/or the state's crime data base. Ten of the 136 men had missing MCMI scales.

A separate regression analysis was conducted for each trait. The MCMI scale score or diagnosis (above 75 BR) and the treatment condition were entered first into logistic regressions, followed by the MCMI-by-treatment type interaction term. A significant increase in R-squared indicates a significant interaction effect. The two traits were also combined with closely related traits through factor analysis and factor scores were used in the equations. Dependent personality was closely linked with somataform disorder. Antisocial personality was closely linked with drug/alcohol abuse potential and histrionic personality.

Table 5 shows the results of the interaction analysis. A diagnosis of dependent personality interacted significantly with treatment. As predicted, those with this disorder had lower recidivism rates for the PPT treatment and higher recidivism for the FCBT treatment. The use of scale scores or factor scores did not produce significant interactions.

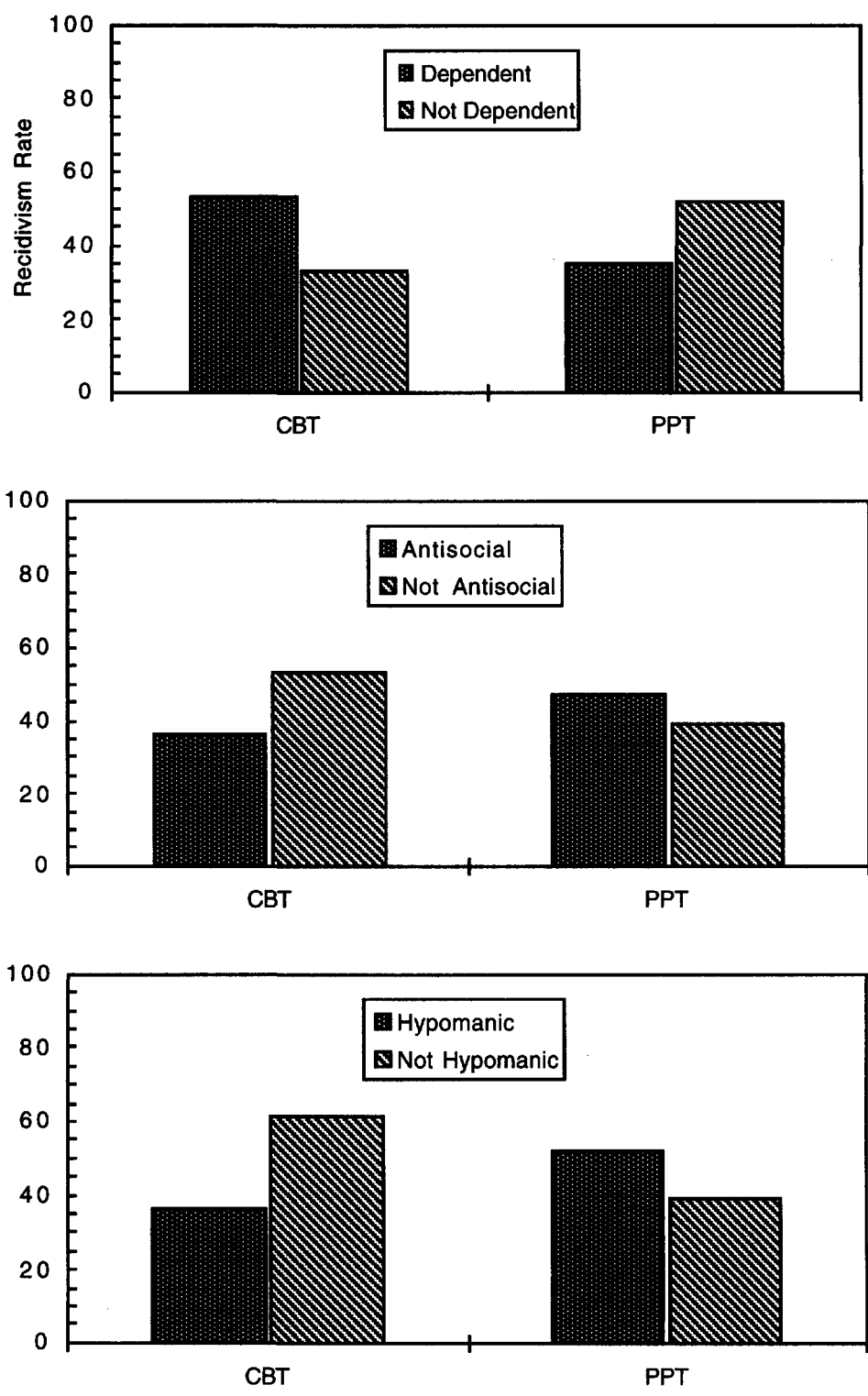
A diagnosis of antisocial personality did not interact significantly with treatment approach, although the relationship was in the predicted direction. However, the scale score and factor score (antisocial with substance abuse potential and histrionic scale scores) did interact significantly with treatment type. Those scoring higher on antisocial personality had lower recidivism rates in FCBT groups and higher rates in the PPT groups.

**TABLE 5. Interaction of Offender Traits  
with Treatment Models in Predicting Recidivism:  
Percentage of Variance Increase from Interaction Term**

| Recidivism: Any Report ( <i>n</i> = 126)     |                      |          |
|----------------------------------------------|----------------------|----------|
|                                              | R <sup>2</sup> incr. | <i>F</i> |
| Dependent Personality Diagnosis              | 2.9%                 | 3.62*    |
| Dependent Personality Score                  | 1.2%                 | 1.43     |
| Dependent/Somataform Factor Score            | 0.8%                 | 1.03     |
| Antisocial Personality Diagnosis             | 1.6%                 | 1.95     |
| Antisocial Personality Score                 | 3.2%                 | 3.94*    |
| Drug/Alc./Histrionic/Antisocial Factor Score | 3.8%                 | 4.95*    |
| Hypomanic Score                              | 4.8%                 | 6.21*    |
| Relationship Satisfaction Score              | 3.7%                 | 4.44*    |

\**p* < .05

**Figure 1.** Interaction of offender traits and type of treatment: Percent with violence recidivism by any report



*Note.* CBT=cognitive-behavioral treatment; PPT=process-psychodynamic treatment

Figure 1 illustrates some of the interactions. Those with a dependent diagnosis had a recidivism rate of 52% if in an FCBT group and only 33% in a PPT group. The opposite was true if they did not have this diagnosis, with a recidivism rate of 35% in an FCBT group and 51% in a PPT group. Those with an antisocial diagnosis had a recidivism rate of 36% if in a FCBT group and 53% if in a PPT group; without an antisocial diagnosis, they had a 47% recidivism rate in the FCBT groups and 39% in the PPT groups.

Two other interactions were significant. Those scoring higher on the hypomanic scale of the MCMI tended to have lower recidivism rates if they were in FCBT groups but higher ones if they were in PPT groups. Those scoring high on this scale tend to be labile, restless, distractible, impulsive, and irritable. If a man had a high BR score (over 65) on this scale, his recidivism rate in the FCBT condition was only 33% and it was 58% in the PPT condition; conversely, a non-hypomanic score led to a 53% recidivism rate in FCBT and a 40% recidivism rate in PPT (Fig 1). Finally, those who were more satisfied with their relationships prior to treatment had lower recidivism rates if they were in the FCBT groups. They had higher recidivism rates if they were in a PPT group.

### Correlates of Personality Traits

Childhood traumas (witnessing or being abused), attitudes, mood states (jealousy, anger, depression), and criminal behavior did not interact significantly with the treatments in predicting outcome. However, many of these variables were significantly related to personality traits in expected directions. There is space here to report only some of the findings. Witnessing parental abuse was significantly related to violence against strangers ( $r = .22, p < .01$ ) and parents ( $r = .24, p < .01$ ). Violence against strangers, in turn, was positively related to antisocial personality ( $r = .21, p < .05$ ) and negatively related to dependent personality ( $r = -.22, p < .01$ ). Severe child abuse was related to violence against parents ( $r = .20, p < .05$ ), which in turn was related to antisocial personality ( $r = .25, p < .01$ ). Attitudes and mood states were more strongly related to a third personality constellation of avoidant, borderline, and depressive traits. Thus, while childhood traumas appear to affect adult behavior and personality, they do not seem to have a direct link to differential outcomes.

### DISCUSSION

This study of the relative efficacy of two treatments for men who batter demonstrates that it is possible to conduct long-term follow-up with a fairly high response rate and to apply treatments in accord with the theoretical orientations espoused by each treatment model. Similar to many studies in other fields that compare "behavioral" and "insight" approaches, there were no differences found between the cognitive-behavioral and process-psychodynamic conditions. Multiple measures and multiple sources of reports were used, relying primarily on the reports of the men's partners 18 to 54 months after treatment. There were no differences reported between treatments in rates of physical abuse or the women's fear levels, general perceptions of change in their partners, or ways of resolving conflicts. These results contrast somewhat with those of Edleson and Syers (1991, 1992) that seemed to favor more structured approaches. The accumulated findings from many studies will obviously be needed to answer questions about the optimal level of structure.

The recidivism rates in this study are in the middle of the range of rates from other abuser outcome studies (Tolman & Edleson, 1995). However, they rose considerably when adding arrest records and show the utility of obtaining these reports. In some cases a woman had not been in contact with her partner, yet an arrest for a crime against a person appearing on his record was most likely from his abuse of another woman.

Because of treatment attrition, the initial randomization to treatments was compromised. The comparison of treatment completers on many behavioral, attitude, demographic, and personality variables on pretreatment measures indicated equivalence. However, it is impossible to know if a key variable went untested. As in many field experiments, the experimental design becomes a quasiexperimental one and only statistical means can be used to try to assure pretreatment equivalence.

The treatment attrition rates were comparable to those of other studies. However, the rate was somewhat lower in the process-psychodynamic treatment (PPT). These results appear to be consistent with those of Stosny (1994) who used a video tape and group discussion early in treatment to arouse the men's compassion to their own traumatic childhoods. Treatment involvement and retention were higher in the "compassion" model. Many programs delay (or never) discuss childhood issues until the final phases of treatment, following accountability, skills training, or other phases. This study suggests that some men may need to work on their childhood traumas early in treatment. The program studied here may have been different than most, however, since considerable work on helping the men increase their accountability for their behavior occurred in individual assessment interviews.

The major finding of this study is that personality styles and disorders interacted with the type of treatment being received. Men with antisocial traits were less likely to be violent after treatment if they attended the feminist-cognitive-behavioral treatment. Men with dependent traits, on the other hand, had better outcomes with the process-psychodynamic treatment. Those with substance abuse potential and hypomania also had lower recidivism in the feminist-cognitive-behavioral condition. The antisocial, substance-abusing offender may need the structure of the FCBT groups. He is more likely to have been severely physically abused in childhood and probably learned to repress most feelings and developed a detached style of relating. The skills-training of the FCBT groups may have matched his action-oriented learning style or his need for structure. The hypomanic offender, sharing the impulsivity of many antisocial offenders, may have benefited from the relaxation training and cognitive restructuring of FCBT groups. The dependent personality probably experienced parental rejection more than direct physical abuse (Dutton, 1994) and developed an anxious attachment style. The unstructured nature of the PPT groups, focusing on group relationships, probably matched the needs of this offender.

Those reporting more satisfying relationships had better outcomes in the FCBT groups. This finding is more difficult to interpret. FCBT may have been more relevant to them because it focused heavily on communication skills and these skills may have been successfully transferred to the home.

There are several important limitations to keep in mind about the findings: (1) A no-treatment control group was not used and thus any reports of change cannot be attributed conclusively to the treatments. Other events in the men's lives, such as arrest or the threat of divorce, may produce substantial change. Although unlikely, the brief interventions that occurred during intake and orientation sessions might also account for lack of group differences. (2) Despite multiple reports of recidivism, the findings are probably underestimates. Reliance on the men's 12-month follow-up reports and on arrest reports are quite likely to be underestimates. Some women may also have underreported their abuse out of shame, fear, or repression of traumatic events. Because the research project was based in

the treatment program, some women may not have trusted the assurances that their reports would be kept from their partners. (3) Some of the measures used with the women were only a single item and have unknown reliability and validity. (4) Despite using experienced group leaders and experienced supervisors, there were no tests of leader competence and thus no assurance the competence levels were equal between the treatments or comparable to other programs. (5) Finally, some of the tests of statistical interaction were post-hoc and need to be replicated in future studies.

In spite of these limitations, this study helps to answer some important methodological questions about the ability to conduct posttreatment follow-up evaluations and test treatment integrity. More important, it is the first step in guiding future studies that would a priori match abuser types with specific kinds of treatment in order to improve outcomes. Progress is being made in such matching in the alcoholism field (e.g., Litt, Babor, DelBoca, Kadden & Cooney, 1992). For practitioners, this study suggests that in developing programs for men who batter, "one size does not fit all." The assumption that all offenders will benefit from highly structured psychoeducational groups that avoid discussion of childhood issues needs to be questioned. Finally, while the results contain signs of hope about treatment effectiveness, a substantial number of men repeated their violence after treatment. More research is needed to identify these treatment failures and to create effective dispositions. These could range from: longer treatment, treatment combined with close probationary supervision or treatment while incarcerated.

## REFERENCES

- Adams, D. (1988). Treatment models of men who batter: A profeminist analysis. In K. Yllo & M. Bograd (Eds.), *Feminist perspectives on wife abuse*. Newbury Park, CA: Sage.
- Andrews, D.A., Zinger, I., Hoge, R.D., Bonta, J., Genreau, P., Cullen, F.T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, 28, 369-404.
- Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 53-63.
- Blood, R.E., & Wolfe, D.M. (1960). *Husbands and wives*. Glencoe, IL: Free Press.
- Browne, K., Saunders, D.G., & Staecker, K.M. (1997). Process-psychodynamic groups for men who batter: Description of a brief treatment model. *Families in Society*, 78, 265-271.
- Buss, A. H., & Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Clinical and Consulting Psychology*, 21, 343-349.
- Chang, H., & Saunders, D.G. (1993). An analysis of attrition in batterers' counseling. Paper presented at the Ninth Annual Symposium on Empirical Group Work, Ann Arbor, Michigan, May, 1993.
- Cook, T.D., & Campbell, D.T. (1979). *Quasi-experimentation: Design and analysis issues in field settings*. Boston: Houghton-Mifflin.
- Deffenbacher, J.L., McNamara, K., Stark, R.S., & Sabadell, P.M. (1990). A comparison of cognitive-behavioral and process-oriented group counseling for general anger reduction. *Journal of Counseling and Development*, 69, 167-172.
- Dunford, F. (1992). The measurement of recidivism in cases of spouse assault. *Journal of Criminal Law and Criminology*, 83, 183-206.

- Dutton, D. (1986). The outcome of court-mandated treatment for wife assault: A quasi-experimental evaluation. *Violence and Victims, 1*, 163-176.
- Dutton, D. (1994). The origin and structure of the abusive personality. *Journal of Personality Disorders, 8* (3), 181-191.
- Dutton, D. (1995). Trauma symptoms and PTSD-Like profiles in perpetrators of intimate abuse. *Journal of Traumatic Stress, 8* (2), 299-316.
- Dutton, D.G., & Starzomski, A.J. (1993). Borderline personality in perpetrators of psychological and physical abuse. *Violence and Victims, 8* (4), 327-337.
- Dutton, D.G., van Ginkel, C., & Starzomski, C. (1995). The role of shame and guilt in the intergenerational transmission of abusiveness. *Violence and Victims, 10*, 121-132.
- Edleson, J.L. (1984). Working with men who batter. *Social Work, 29*, 237-242.
- Edleson, J.L. & Syers, M. (1990). The relative effectiveness of group treatment for men who batter. *Social Work Research and Abstracts, 26*, 10-17.
- Edleson, J. L., & Syers, M. (1991). The effects of group treatment for men who batter: An 18 month follow-up study. *Research on Social Work Practice, 1*, 227-243.
- Edleson, J.L., & Tolman, R.M. (1992). *Intervention for men who batter: An ecological approach*. Newbury Park, CA: Sage.
- Fleming, & Courtney, (1984). The dimensionality of self-esteem: II. Hierarchical facet model for revised measurement scales. *Journal of Personality and Social Psychology, 46*, 404-421.
- Ganley, A. L. (1989). Integrating feminist and social learning analyses of aggression. In P. L. Caesar & L. K. Hamberger (Eds.), *Treating men who batter*. New York: Springer Publishing Company.
- Gil, E. (1983). *Outgrowing the pain*. New York: Dell.
- Greenwald, H.J., & Satow, Y. (1970). A short social desirability scale. *Psychological Reports, 27*, 131-135.
- Hamberger, L.K., & Hastings, J.E. (1986). Personality correlates of men who abuse their partners: A cross-validation study. *Journal of Family Violence, 1* (4), 323-341.
- Hamberger, L. K., & Hastings, J. E. (1993). Court-mandated treatment of men who assault their partners: Issues, controversies, and outcomes. In N. Z. Hilton (Ed.), *Legal responses to domestic violence*. Newbury Park, CA: Sage.
- Harrell, A. V. (1991). Evaluation of court-ordered treatment for domestic violence. Final report submitted to the State Justice Institute. Washington, DC: The Urban Institute.
- Holtzworth-Munroe, A., & Stuart, G.L. (1994). Typologies of male batterers: Three subtypes and the differences among them. *Psychological Bulletin, 116* (3), 476-497.
- Hotaling, G. T., & Sugarman, D. B. (1986). An analysis of risk markers in husband to wife violence: The current state of knowledge. *Violence and Victims, 1*, 101-124.
- Jennings, J. L. (1987). History and issues in the treatment of battering men: A case for unstructured group therapy. *Journal of Family Violence, 2*, 193-214.
- Jones, E. E., & Pulos, S. M. (1993). Comparing the process in psychodynamic and cognitive-behavioral therapies. *Journal of Consulting and Clinical Psychology, 61*, 306-316.
- Litt, M.D., Babor, T.F., DelBoca, F.K., Kadden, M., & Cooney, N.L. (1992). Types of alcoholic, II: Application of an empirically derived typology to treatment matching. *Archives of General Psychiatry, 49*, 609-614.
- Millon, T. (1983). *Millon Clinical Multiaxial Inventory*. Minneapolis, MN: National Computer Systems.
- Novaco, R. W. (1975). *Anger control*. Lexington, MA: Lexington.
- O'Leary, K.D. (1988). Physical aggression between spouses: A social learning perspective. In van Hasselt, Morrison, Bellak, & Hersen (Eds.), *Handbook of family violence*. New York: Plenum.

- Parker, B., & Ulrich, Y. (1990). A protocol of safety: Research on abuse of women. *Nursing Research*, 39, 248-250.
- Roach, A.J., Frazier, L.P., & Bowden, S.R. (1981). The marital satisfaction scale: Development of a measure for intervention research. *Journal of Marriage and the Family*, 9, 130-141.
- Rosenfeld, B. D. (1992). Court-ordered treatment of spouse abuse. *Clinical Psychology Review*, 12, 205-226.
- Rumptz, M.H., Sullivan, C.M., Davidson, W.S., & Basta, J. (1991). An ecological approach to tracking battered women over time. *Violence and Victims*, 6, 237-244.
- Saunders, D.G. (1984). Helping husbands who batter. *Social Casework*, 65(6), 347-353.
- Saunders, D.G., & Azar, S. (1989). Treatment programs for family violence. In L. Ohlin & M. Tonry (Eds.), *Family violence: Crime and justice, a review of research (Vol. II)*. Chicago, IL: University of Chicago Press.
- Saunders, D. G., & Hanusa, D. R. (1986). Cognitive-behavioral treatment of men who batter: The short-term effect of group therapy. *Journal of Family Violence*, 1(4), 357-372.
- Saunders, D. G., Lynch, A. E., Grayson, M., & Linz, D. (1987). The inventory of beliefs about wife beating. The development and initial validation of a measure of beliefs and attitudes. *Violence and Victims*, 2, 39-57.
- Scher, M., & Stevens, M. (1987). Men and violence. *Journal of Counseling and Development*, 65, 351-355.
- Sonkin, D. J. (1989). *Learning to live without violence*. CA: Volcano Press.
- Spence, J.T. & Helmreich, R.L. (1978). *Masculinity and femininity: Their psychological dimensions, correlates, and antecedents*. University of Texas, Austin.
- Stordeur, R. & Stilles, R. (1989). *Ending men's violence against their partners*. Newbury Park, CA: Sage.
- Stosny, S. (1994). Shadows of the Heart: A dramatic video for the treatment resistance of spouse abusers. *Social Work*, 39, 686-694.
- Straus, M.A. & Gelles, R.J. (1990) *Physical violence in American families*. New Brunswick, NJ: Transaction.
- Straus, M. A., Gelles, R. J., & Steinmetz, S. K. (1980). *Behind closed doors: Violence in the American family*. New York: Doubleday/Anchor.
- Tolman, R. M., & Bennett, L. W. (1990). A review of research on men who batter. *Journal of Interpersonal Violence*, 5, 87-118.
- Tolman, R. M., & Edleson, J. L. (1995). Intervention for men who batter: A review of research. In S. Stith & M. A. Straus (Eds.), *Understanding intimate violence*. Minneapolis: National Council on Family Relations.
- van Hasselt, V.B., Morrison, R.L., Bellack, A.S., & Hersen, M. (1988). *Handbook of family violence*. New York: Plenum.
- White, G.L. (1977). The social psychology of romantic jealousy. *Dissertation Abstracts International*, 37(10), 5449-B.
- Whitfield, C.L. (1987). *Healing the Inner Child*. Deerfield Beach, FL: Health Communications.
- Yllo, K. (1984). Patriarchy and violence against wives: The impact of structural and normative factors. *Journal of International and Comparative Social Welfare*, 1, 16-29.

*Acknowledgements* Funding for this study was provided by the Centers for Disease Control, Grant Award No. R49/CCR 502584-03, "Preventing Injuries from Marital Violence" to Daniel G. Saunders, Principal Investigator; the Research and Development Fund, Department of Psychiatry, University of Wisconsin-Madison, 133-P978; and the Office of the Vice President for Research,



University of Michigan. The author is grateful to the sponsoring program, Alternatives to Aggression, Family Service, Madison, Wisconsin, and to all those who worked on the project, including the consultants, group leaders, interviewers, and office workers. He is especially grateful to the project coordinators, Patricia Cumbie, Susan Shepanek, Judith Witt, and Randy Magen, and to the group supervisors, Kevin Browne, Darald Hanusa, Dorothy Helman., Kathryn Staecker, and Pauline Thome.

*Offprints.* Requests for offprints should be directed to Daniel G. Saunders, School of Social Work, University of Michigan, 1065 Frieze Building, Ann Arbor, MI 48109-1285.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

# Interventions for Intimate Partner Violence: Review and Implications for Evidence-Based Practice

Carla Smith Stover

Yale University and the National Center for Children Exposed to Violence

Amy Lynn Meadows and Joan Kaufman

Yale University

The objective of this article was to survey available intimate partner violence (IPV) treatment studies with (a) randomized case assignment, and (b) at least 20 participants per group. Studies were classified into 4 categories according to primary treatment focus: perpetrator, victim, couples, or child-witness interventions. The results suggest that extant interventions have limited effect on repeat violence, with most treatments reporting minimal benefit above arrest alone. There is a lack of research evidence for the effectiveness of the most common treatments provided for victims and perpetrators of IPV, including the Duluth model for perpetrators and shelter–advocacy approaches for victims. Rates of recidivism in most perpetrator- and partner-focused treatments are approximately 30% within 6 months, regardless of intervention strategy used. Couples treatment approaches that simultaneously address problems with substance abuse and aggression yield the lowest recidivism rates, and manualized child trauma treatments are effective in reducing child symptoms secondary to IPV. This review shows the benefit of integrating empirically validated substance abuse and trauma treatments into IPV interventions and highlights the need for more work in this area.

**Keywords:** intimate partner violence, treatment, batterers, child witness to violence

Intimate partner violence (IPV) impacts millions of families worldwide (Watts & Zimmerman, 2002). In the United States alone, lifetime prevalence studies suggest between 20% and 30% of women will be assaulted by an intimate partner and between 5%

and 20% of children will witness a parent being assaulted (McCloskey & Walker, 2000; Tjaden & Thoennes, 2000; Wilt & Olson, 1996). The impact of IPV is well documented in the research literature, with deleterious effects acknowledged for all members of the family (Carter, Weithorn, & Behrman, 1999). For several decades, law enforcement, courts, social service agencies, and mental health providers have attempted to develop interventions to assist victims of IPV and prevent batterers from continuing to use violence in their relationships.

Reducing violence perpetration has proven a challenge, however, as perpetrators of IPV have complicated psychosocial and psychiatric histories. Many have witnessed family violence or were victims of abuse as children (Gortner, Gollan, & Jacobson, 1997). In addition, borderline, narcissistic, and antisocial personality disorders are common among IPV perpetrators (Mauricio, Tein, & Lopez, 2007), and the co-occurrence of substance abuse problems in this population is high, with rates ranging from 40% to 92% across studies (Brookoff, O'Brien, Cook, Thompson, & Williams, 1997; Easton, Swan, & Sinha, 2000; Wilt & Olson, 1996).

Despite the frequent co-occurrence of these problems, incorporation of the perpetrator's own trauma history, personality disorders, and substance abuse are not typically targeted into IPV intervention models. Some of the first studies to evaluate strategies for IPV assessed the impact of mandatory arrest, which required officers to make an arrest or issue a warrant for the perpetrator of violence at the time of the incident in every case of IPV (Sherman & Berk, 1984). This policy eliminated officer discretion in determining the need for an arrest. It was thought that the criminal justice ramifications would deter perpetrators from continuing to use violence.

Aside from mandatory arrest, the standard for batterers' intervention is a group treatment that focuses on feminist psychoedu-

---

*Editor's Note.* This is one of six accepted articles received in response to an open call for submissions on interventions for intimate partner violence/domestic violence.—MCR

---

CARLA SMITH STOVER received her PhD in clinical psychology from Alliant International University—San Francisco. She is an associate research scientist and the research coordinator for the National Center for Children Exposed to Violence at the Yale Child Study Center. Her areas of research include development and evaluation of treatment for children and families exposed to violence and trauma.

AMY LYNN MEADOWS received her MD from Yale University School of Medicine. She soon will begin a triple board program in pediatrics, psychiatry, and child psychiatry at University of Pittsburgh Medical Center.

JOAN KAUFMAN received her PhD in clinical psychology from Yale University. She is currently an associate professor within the Department of Psychiatry at Yale University School of Medicine. She is director of Yale University's Child and Adolescent Research and Education (CARE) Program, which is dedicated to working with maltreated children and their families and has foci that span from neurobiology to social policy.

THE WORK FOR THIS ARTICLE WAS SUPPORTED by the National Center for Children Exposed to Violence (for Carla Smith Stover) and by the National Center for Posttraumatic Stress Disorder, Veterans Affairs New England Healthcare System, West Haven, Connecticut (for Joan Kaufman).

CORRESPONDENCE CONCERNING THIS ARTICLE should be addressed to Carla Smith Stover, Yale University Child Study Center, 230 S. Frontage Rd., New Haven, CT 06520. E-mail: carla.stover@yale.edu

cation about power and control often referred to as the *Duluth model* (Pence & Paymar, 1993). According to this model, the primary cause of domestic violence is patriarchal ideology and societal sanctioning of men's power and control over women. The fundamental tool of the Duluth model is the *Power and Control Wheel*, which illustrates how men use intimidation, male privilege, isolation, emotional and economic abuse, and violence to control women. The model is implemented in a variety of protocols, lasting 8–36 weeks, and is the unchallenged treatment of choice in most communities. In some states, it is the mandated treatment.

Another common approach to batterer treatment is group cognitive-behavioral treatment (CBT), in which learning nonviolence is the primary focus (Adams, 1988). The CBT therapist works to point out the pros and cons of violence, along with providing skills training (e.g., anger management, communication skills, assertiveness, relaxation techniques) to promote alternatives to violence. Programs have also combined aspects of both the Duluth and CBT models, and distinguishing between the two is becoming increasingly difficult.

In addition to focusing on the needs of perpetrators, numerous IPV interventions aim to address the needs of their partners. Partners of batterers are at risk for a range of negative consequences that go beyond immediate physical injuries to include a variety of stress-related psychiatric disorders (Campbell et al., 2002; Eisenstat & Bancroft, 1999). Associated psychiatric symptoms can be profound and include depression, posttraumatic stress disorder (PTSD), and other anxiety disorders (Dutton et al., 2006; Golding, 1999). Treatments for victims of IPV typically focus on advocacy and counseling to assist the victims in leaving their abusive partners, with the most commonly evaluated services provided by domestic violence shelters.

Child-focused interventions aim to address the most common sequelae experienced by children exposed to domestic violence. Child witnesses are at increased risk for attachment disorders, depression, PTSD, other anxiety disorders, and conduct problems (e.g., Kendall-Tackett, 2004). Standard care for child witnesses involves group treatment while in shelter with their mothers or referral for individual treatment within a community mental health clinic. More recently, several treatment approaches for child witnesses of IPV have been manualized and published (Cohen, Mannarino, & Deblinger, 2006; Lieberman & Van Horn, 2004).

Two recent meta-analytic studies evaluated the efficacy of batterers' treatment programs, but the vast majority of studies included were quasi-experimental (Babcock, Green, & Robie, 2004; Feder & Wilson, 2005). This review focuses on randomized controlled studies and expands the review of empirical research on IPV treatments to include interventions developed for IPV partners and their children. The rationale for focusing on treatments targeting perpetrators, partners, and their children is derived from the frequent co-occurrence of IPV and child abuse and the common practice of protective services workers to mandate treatment for all members of the family when domestic violence is a presenting issue. In addition, as noted earlier, it has been well documented that IPV has deleterious effects on all members of the family, further highlighting the need for efficacious treatments for perpetrators, their partners, and child witnesses of IPV.

## Method

MEDLINE and PsycINFO data bases were searched from their respective start dates to June 2007 using specific keywords such as *domestic violence*, *batterers*, *partner abuse*, *intimate partner violence*, *domestic violence intervention*, *children and domestic violence*, *batterer treatment*, and *domestic violence couples treatment*.

Bibliographies of key articles were searched by hand.

Identified IPV interventions were categorized into the relevant client categories (batterer, victim, couple, and child witness). This search yielded 30 batterer, 18 victim, 18 couples, and 19 child-witness intervention evaluations. Studies included in this review met the following criteria: (a) experimental study (randomized treatment and control), (b) sample size of at least 20 participants per group, and (c) recidivism or measures of violence severity as an outcome variable. Application of these selection criteria, however, resulted in identification of only one couple and no child-witness treatment studies. Given that only one couple treatment evaluation utilized a randomized control group, we also included studies that compared couple therapy with another treatment modality for IPV in this review. In addition, relaxing the last criterion of recidivism as an outcome variable resulted in the identification of four evaluation studies that assessed change in symptoms in interventions targeting child witnesses of IPV. In total, seven batterer, six victim, five couple, and four child-witness treatment studies were surveyed in this review. We will describe novel, promising interventions more fully.

## Results

### *Interventions for Batterers*

The treatments for perpetrators reviewed in this section are summarized in Table 1. The preponderance of research examined the effect of mandatory arrest or group treatment models. As can be seen in the table, participant dropout was a significant problem for group treatment approaches, with rates approximately 30% across studies. Attrition was also a significant problem in most studies in which recidivism rates relied on victim response, with loss to follow-up rates ranging from 15% to 89%. In addition, when recidivism rates were calculated from police report and victim response, rates were consistently and notably higher when based on victim report. Given the high rate of victim data missing in most studies, the reported recidivism rates should be accepted with caution.

*Mandatory arrest.* In an initial study of mandatory arrest in Minneapolis, Minnesota, 314 cases of simple (misdemeanor) assault were randomly assigned to receive one of three responses: mandatory arrest of the perpetrator, mediation by the responding officer, or physical separation of the couple for 8 hr. Of suspects randomized to arrest, 99% were arrested, but only approximately three fourths of the subjects in the other conditions received the intervention they were assigned. On the basis of the 12-month follow-up police record data, mandatory arrest resulted in a 13% recidivism rate compared with 26% for those separated from their partners for 8 hr. The recidivism rate for those who received mediation fell midway between and was statistically indistinguishable from the other two groups. Only 49% of the victims were reached for 12-month follow-up, with reported victim recidivism rates of 19% for mandatory arrest and 37% for mediation condi-

tions; recidivism rates for those in the separation condition were midway between and again statistically indistinguishable from the other two groups (Sherman & Berk, 1984).

A large-scale ( $N = 4,032$ ) multisite replication and analysis of mandatory arrest for domestic violence failed to demonstrate a benefit of mandatory arrest on perpetrator violence on the basis of police report data (Spousal Assault Replication Project, or SARP; Maxwell, Garner, & Fagan, 2001). A significantly lower rate of recidivism was reported for arrest on the basis of victim report data, but high rates were still present in both mandatory arrest cases and controls (36% vs. 48%).

*Duluth model of group treatment.* Two studies evaluated batterers treated with the Duluth model compared with a control group. In the first study, men were randomly assigned either to a 26-week Duluth model group plus probation or to probation only; both interventions were associated with a 24% recidivism rate (Feder & Dugan, 2002). Treatment completers were less likely to be rearrested (13%) compared with noncompleters (30%). In the second study, men were randomly assigned either to 40 hr of Duluth model group treatment (in either 26-week or 8-week format) or to a community service control (40 hr of service completed in a 2-week period). Men randomized in this study agreed to engage in treatment as part of their sentence, biasing the sample to more treatment-motivated men. Recidivism rates were 16% and 26%, respectively, according to police report, and 22% and 21%, respectively, according to victim report (Taylor, Davis, & Maxwell, 2001). Attrition rates were high, with only 30%–50% of victims responding at 12-month follow-up.

*Group CBT or combined CBT–psychoeducation interventions.* Dunford (2000) conducted the most methodologically rigorous study to date comparing CBT men's groups with conjoint couple therapy groups and no treatment controls. According to police or victim reports, neither treatment had a significant impact on recidivism for this sample of military men at 1-year follow-up. Rates of police-reported recidivism were extremely low in this sample (3%–6%), and consistent with other studies, victim reports yielded considerably higher rates of repeat violence (range: 27%–35%, no difference between groups).

Palmer, Brown, and Barrera (1992) randomly assigned 56 Canadian men to either a 10-week group treatment (combined CBT and psychoeducation) or a no-treatment control group. Based on police records, recidivism rates were significantly higher for controls (31%) than for the intervention group (10%). This study had a small sample size with only 22% of victims responding at 12-month follow-up.

Ford and Regoli (1993) randomly assigned 347 men to pretrial counseling (type of counseling not specified), counseling as probation, or mandatory sentencing. They found that pretrial counseling was more effective than counseling as a condition of probation (recidivism rate: 34% vs. 45%) but no more effective than mandatory sentencing (recidivism rate: 34%). Only 31% of victims were reached for 6-month follow-up assessments.

*Summary of batterer treatments.* Group treatments for IPV batterers have meager effects on the cycle of violence, with most studies demonstrating no or minimal impact above that of mandatory arrest alone. Most studies, regardless of intervention strategy (mandatory arrest, Duluth model group treatment, CBT), report approximately one in three cases will have a new episode of IPV within 6 months based on victim's reports. This rate must be accepted with caution

given high attrition in victim reports across studies (range: 15%–78%; mean attrition: 46%).

### *Interventions for Victims of IPV*

Interventions that have been evaluated for victims of IPV have been based in (a) shelters, (b) prenatal clinics, or (c) the community, with police–social service outreach and advocacy (see Table 2). These studies had significantly lower follow-up attrition rates than the interventions targeting perpetrators but have reported recidivism rates comparable to, or greater than, those reported in perpetrator-focused studies.

*Shelter interventions.* In the only methodologically sound set of studies evaluating an intervention for victims exiting shelter, Sullivan and colleagues examined the efficacy of a 10-week advocacy program for women after at least 1 night's stay in a domestic violence shelter. The program included 4–6 hr per week of one-on-one advocacy and counseling. The initial sample of 141 participants did not experience significant differences in repeat violence at 6-month follow-up (Sullivan, Campbell, Angelique, Eby, & Davidson, 1994). Further data collection in a total sample of 278 women interviewed every 6 months for 2 years and in a subset of 124 women reinterviewed at 3 years revealed a modest reduction in revictimization rates between 6 months and 2 year postintervention (31% intervention vs. 37% controls; Sullivan & Bybee, 1999). However, these differences were not sustained for the subsample followed through 3 years (44% intervention vs. 36% controls; Bybee & Sullivan, 2005).

*Prenatal clinic interventions.* In a study with 329 Hispanic victims of IPV seen in a prenatal clinic, McFarlane, Soeken, and Wiist (2000) compared three interventions: (a) wallet-sized resource cards, (b) unlimited access to supportive, nondirective counseling, or (c) unlimited counseling plus support from a “mentor mother.” Interventions were provided during the prenatal period only, and women were interviewed at 2, 6, 12, and 18 months postdelivery. Although women who received both counseling and mentorship reported less violence at 2 months postdelivery than did the counseling-only group, so did the resource-card group. There were no significant differences among the groups at 12 or 18 months postdelivery. The potency of these intervention strategies cannot be fully evaluated, however, as exact recidivism rates were not reported in the study.

*Police–social service outreach programs.* Several police–social service outreach programs have been developed in various communities. The Domestic Violence Intervention Education Project (DVIEP) was conducted in the New York City public housing projects. The DVIEP involved follow-up home visits made by police officers and social workers to homes where a domestic dispute was reported to the police to provide victims with information on services available to them. Results of the study indicated that victims who received the DVIEP were more likely to call the police and to call more rapidly to report abuse in the 6 months following the intervention than those assigned to the comparison group (45% vs. 39%, respectively; Davis et al., 2003). However, on the basis of victim report from 72% of the sample at 6-month follow-up, there were no group differences in severity of abuse reported on the Conflict Tactics Scale, with high rates of recidivism reported in both groups

Table 1  
*Interventions for Batterers*

| Citation                       | Sample <i>N</i> | Treatment                                                                                                   | Recidivism measure                                     | Posttreatment follow-up | Significance outcome                                                                                         | Recidivism by group                                                                                                                                                                    | Treatment dropout | Follow-up attrition                            |
|--------------------------------|-----------------|-------------------------------------------------------------------------------------------------------------|--------------------------------------------------------|-------------------------|--------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|------------------------------------------------|
| Sherman & Berk, 1984           | 314             | Mandatory arrest (136) vs. ordering offender from premises for 8 hr (89) vs. informal mediation (89)        | Police and victim report                               | 6 months                | Mandatory arrest significantly more effective deterrent than either physical separation or officer mediation | 26% separation vs. 13% arrest based on police record; 37% mediation vs. 19% arrest based on victim report; mediation statistically indistinguishable from both groups on both measures | Not applicable    | 0% for police record; 51% for victim interview |
| Maxwell, Garner, & Fagan, 2001 | 4,032           | Mandatory arrest (1,748) vs. nonarrest (2,284)                                                              | Police and victim report                               | 6 months                | No differences between groups on police report; arrest = significantly less violence based on victim report  | 36% arrest vs. 48% nonarrest based on victim report                                                                                                                                    | Not applicable    | 0% for police record; 37% for victim report    |
| Feder & Dugan, 2002            | 404             | 26-week Duluth model tx group + probation (216) vs. probation only (188)                                    | Partner and self-report CTS, police report             | 12 months               | No significant difference in rates of rearrest, attitudes, or incidence of violence                          | 24% in both groups rearrested in 1 year; 30% tx noncompleters rearrested vs. 13% completers                                                                                            | 29%               | 70% for victim report at 12 months             |
| Taylor, Davis, & Maxwell, 2001 | 376             | 40 hr of Duluth model tx group (190) vs. 40 hr of community service control (186)                           | Partner report CTS, police report                      | 12 months               | 26 wk lower recidivism at 6 and 12 mo; no difference for victim report of violence                           | 16% Duluth model tx vs. 26% control based on police record; 22% tx vs. 21% control based on victim report                                                                              | Not reported      | 50% for victim report                          |
| Dunford, 2000                  | 861             | Men's CBT group (168) vs. conjoint group (153) vs. rigorously monitored group (173) vs. control group (150) | Partner and self-report MCTS, police and court records | 18 months               | No differences between groups                                                                                | 4% men's group, 3% couples, 6% monitored, 4% control based on police record; 29% men's group, 30% couples, 27% monitored, 35% control based on victim report                           | 29%               | 0% for police record; 15% for victim report    |

(table continues)



Table 1 (continued)

| Citation                       | Sample <i>N</i> | Treatment                                                                                | Recidivism measure        | Posttreatment follow-up | Significance outcome                                      | Recidivism by group                                                                                              | Treatment dropout | Follow-up attrition                                 |
|--------------------------------|-----------------|------------------------------------------------------------------------------------------|---------------------------|-------------------------|-----------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|-------------------|-----------------------------------------------------|
| Palmer, Brown, & Barrera, 1992 | 56              | 10-week psychoeducation group (30) vs. probation only (26)                               | Partner and police report | 12 months               | Recidivism significantly higher for control vs. treatment | 10% psychoeducation vs. 31% control based on police report; victim data not reported due to high study attrition | 30%               | 70% intervention and 89% control for victim reports |
| Ford & Regoli, 1993            | 347             | Pretrial counseling (127) vs. counseling as probation (114) vs. mandatory sentence (106) | Partner report            | 6 months                | No difference between groups                              | 34% pretrial vs. 45% probation vs. 34% mandatory sentence                                                        | Notreported       | 37% for victim report at 6 months                   |

Note. Tx = treatment; CTS = Conflict Tactics Scale; CBT = cognitive-behavioral therapy; MCTS = Modified Conflict Tactics Scale.

(45% DVIEP vs. 39% control). These results are consistent with the evaluation of another nonrandomized study that compared a similar police–advocacy intervention that was provided in five police districts with a no-treatment control group of IPV cases in five comparison police districts (Stover, Berkman, Desai, & Marans, 2008).

*Summary of victim treatments.* Studies of victim interventions reveal higher recidivism rates overall than batterer treatment approaches, regardless of whether victim or official police records are used. Rates ranged from 31% to 44%. It appears that postshelter support and advocacy approaches have short-term impacts that are less effective than mandatory arrest, and none of the other approaches examined to date are effective in reducing subsequent violence.

### Couple Treatment for IPV

*Couple-focused interventions.* Couple treatment studies had the least methodological rigor. Only one study utilized a randomized control condition (see Table 3). The four other studies included in the table compared several types of treatments without a control group. Treatment completion and recidivism rates varied considerably from study to study, with no consistent patterning of findings to explain variability in rates across studies.

As detailed earlier, Dunford (2000) found no group differences for couple treatment, men's CBT, or controls in reducing IPV recidivism for active-duty army personnel. Harris, Savage, Jones, & Brooke (1988) randomly assigned 58 couples to either a multicouple group or individual couple counseling. While only 16% of the 23 couples assigned to the multicouple group condition dropped out, 67% of the 35 couples assigned to individual couple counseling dropped out before completing treatment. For treatment completers, no significant differences in recidivism were found between the two treatments. Overall, a 20% recidivism rate was reported at 6-month follow-up, but given the high dropout rate, between-group comparisons could not be made.

O'Leary, Heyman, and Neidig (1999) assigned 75 volunteer couples to either feminist cognitive-behavioral gender-specific groups or conjoint treatment. Dropout rates were high, limiting the ability of the investigators to compare group outcomes. For treatment completers, violence severity ratings had decreased approximately 50% by posttreatment and were comparably low at 1-year follow up. However, recidivism rates were 74% overall, with no between-group recidivism analyses conducted. A second study examining these two modes of treatment with 49 couples reported notably lower dropout and recidivism rates (Brannen & Rubin, 1996). The sample for this latter study was court referred and limited to men with alcohol use disorders.

Fals-Stewart, Kashdan, O'Farrell, and Birchler (2002) found behavioral couples therapy (BCT) was more effective than individual substance abuse treatment in reducing recidivism for men with comorbid substance abuse and domestic violence, with rates of recidivism at 18% for BCT versus 43% for individual treatment at 12-month follow-up. In BCT, men receive weekly individual and group drug abuse counseling (both of which emphasize cognitive-behavioral anger management and coping skills training). Additionally, males and their female partners meet conjointly for weekly BCT sessions. The BCT sessions, which are described in greater detail by O'Farrell and Fals-Stewart (2006),

are used to (a) help male partners remain abstinent, (b) teach more effective communication skills, and (c) increase positive behavioral exchanges between partners (Fals-Stewart et al., 2002). While not initially developed to target IPV, the CBT portion of BCT includes many of the CBT approaches used in batterer programs. The addition of substance abuse and couples treatment foci appears to have contributed significantly to the lower dropout rate and greater reduction in violence for men participating in this intervention.

*Summary of couples treatment.* The studies reviewed in this section provide preliminary data to support the efficacy of BCT and multigroup couples interventions for IPV for perpetrators of violence struggling with alcohol and substance use disorders. The efficacy of these approaches when substance use is not identified or addressed has not been consistently supported.

### *Treatments for Children Exposed to IPV*

*Child-witness interventions.* Studies that measured recidivism as an outcome for child-witness-to-IPV treatments were not found. Instead, four studies were identified that were designed to assess reductions in symptoms of children exposed to violence (see

Table 4). Child-parent psychotherapy (CPP; Lieberman & Van Horn, 2004) was developed to address the needs of preschool children exposed to family violence. It is a 52-week dyadic treatment that integrates modalities derived from psychodynamic, attachment, trauma, cognitive-behavioral, and social learning theories. A randomized controlled trial of CPP for young children exposed to domestic violence resulted in significant reductions in both child and parent symptoms posttreatment and at 6-month follow-up (Lieberman, Ghosh Ippen, & Van Horn, 2006; Lieberman, Van Horn, & Ghosh Ippen, 2005). In their evaluation, however, Lieberman and Van Horn required the mothers to have ended their relationship with the violent partner, have separate stable housing, and have been clean of substances for 6 months. These exclusion criteria did not allow participation of couples who remained together or those struggling with substance abuse.

Another study of children 6–12 years old who were exposed to IPV compared a 10-week group treatment program for children only (CO) with a 10-week program of combined concurrent group sessions for children and their mothers (CM) and a wait-listed control group. The children's groups provided psy-

Table 2  
*Interventions for Victims of Intimate Partner Violence*

| Citation                         | Sample and<br>N                        | Treatment (n)                                                                               | Recidivism<br>measure         | Posttreatment<br>follow-up | Significance outcome                                                                                                                        | Recidivism by<br>group                                                      | Treatment<br>dropout                                          | Follow-up<br>attrition |
|----------------------------------|----------------------------------------|---------------------------------------------------------------------------------------------|-------------------------------|----------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------|---------------------------------------------------------------|------------------------|
| Sullivan et al., 1994            | DV shelter; 141                        | 10 weeks <sup>a</sup> of postshelter advocacy (71) or shelter only (70)                     | Victim report CTS             | 6 months                   | No difference between groups; quality of life and social supports improved in both groups                                                   | 43% experienced further abuse                                               | Not reported                                                  | Not reported           |
| Sullivan & Bybee, 1999           | DV shelter; 278                        | 10 weeks <sup>a</sup> of postshelter advocacy counseling (135) vs. shelter alone (130)      | Victim report CTS             | 2 years                    | Intervention group = less violence and less risk for reabuse, but overall significant decrease for both groups                              | 31% intervention vs. 37% control at 2-year follow-up                        | Not reported                                                  | 5% at 2 years          |
| Bybee & Sullivan, 2005           | DV shelter; 124                        | 10 weeks <sup>a</sup> of postshelter advocacy counseling (71) vs. shelter only (70)         | Victim report CTS             | 3 years                    | No differences between groups                                                                                                               | 44% intervention vs. 36% control between 2 and 3 years                      | Not reported                                                  | 12% at 3 years         |
| McFarlane, Soeken, & Wiist, 2000 | Prenatal clinic; 329                   | Brief <sup>b</sup> (94) vs. counseling <sup>c</sup> (73) vs. lay outreach <sup>d</sup> (92) | Victim report CTS             | 2, 6, 12, and 18 months    | Outreach decreased violence scores at 2 months postdelivery more than counseling alone, but not sustained at 6-, 12-, or 18-month follow-up | No effect of intervention on elimination of abuse; percentage not reported. | Not reported                                                  | 21% at 18 months       |
| Davis, Maxwell, & Taylor, 2006   | IPV cases in NYC housing projects; 434 | DVIEP (police-social worker home visit) vs. control                                         | Police and victim report CTS2 | 6 months                   | Significantly more police calls in DVIEP group, but no difference in CTS2 severity                                                          | 45% DVIEP vs. 39% control based on police report; no victim report provided | All intervention cases received at least one DVIEP home visit | 28%                    |

*Note.* DV = domestic violence; CTS = Conflict Tactics Scale; IPV = intimate partner violence; NYC = New York City; DVIEP = Domestic Violence Intervention Education Project.

<sup>a</sup> 4–6 hr/week. <sup>b</sup> Wallet card with resource information. <sup>c</sup> Unlimited access to DV counselor. <sup>d</sup> Unlimited professional counseling plus “mentor mother.”



Table 3  
*Interventions for Couples in Violent Relationships*

| Citation                  | Sample N                              | Treatment (n)                                                                                               | Outcome measure                                        | Posttreatment follow-up | Significance outcome                                                                                                                                     | Recidivism by group                                | Treatment dropout                                       | Follow-up attrition                 |
|---------------------------|---------------------------------------|-------------------------------------------------------------------------------------------------------------|--------------------------------------------------------|-------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------|---------------------------------------------------------|-------------------------------------|
| Dunford, 2000             | 620 IPV couples who were active army  | Men's CBT group (168) vs. conjoint group (153) vs. rigorously monitored group (173) vs. control group (150) | Partner and self report MCTS, police and court records | 18 months               | No differences among groups                                                                                                                              | 29% men's, 30% couples, 27% monitored, 35% control | 29%                                                     | 0% police report; 15% victim report |
| Harris et al., 1988       | 58 IPV couples in community           | Couples group (23) vs. individual tx (35) vs. wait-listed controls (10) <sup>a</sup>                        | Partner report CTS                                     | 12 months               | No significant difference at 6–12 mo follow-up                                                                                                           | 18% overall                                        | 47% total dropout: 16% couples group, 67% individual tx | 59%                                 |
| O'Leary et al., 1999      | 74 IPV couples                        | 14 sessions of gender-specific group (30 couples) vs. conjoint tx group (44 couples)                        | Partner report MCTS                                    | 12 months               | Improvement in marital adjustment, husbands' taking responsibility for aggression in both tx conditions, with no difference between groups in recidivism | 74% overall                                        | 50% gender specific, 45% conjoint, 47% overall          | 16%                                 |
| Brannen & Rubin, 1996     | 49 IPV couples                        | Gender-specific group (26) vs. conjoint group (22) <sup>b</sup>                                             | Partner report MCTS                                    | 6 months                | In couples with EtOH, conjoint treatment was superior, but gains were not sustained in either at 6-month follow-up                                       | 8.3% gender-specific vs. 7.1% conjoint             | 24%                                                     | 57%                                 |
| Fals-Stewart et al., 2002 | 86 IPV couples with alcoholic husband | BCT (43 couples) vs. individual tx (43 individuals)                                                         | Partner report of violence                             | 12 months               | Lower violence scores in couples tx for pts with comorbid substance abuse                                                                                | 18% BCT vs. 43% individual tx                      | 14%                                                     | Not reported                        |

*Note.* CBT = cognitive-behavioral therapy; MCTS = Modified Conflict Tactics Scale; CTS = Conflict Tactics Scale; BCT = behavioral couples therapy; tx = treatment; EtOH = alcoholism.  
<sup>a</sup> Couples group was 10 sessions (3 hr each) with same-sex peer group and teaching sessions with both genders; individual treatment was family systems based with an open-ended time frame. Wait-listed controls received treatment after 10 weeks. <sup>b</sup> Gender-specific group was 12 weeks (1.5-hr session twice a week) of Domestic Abuse Project model treatment. Conjoint group received CBT with change, anger, and problem-solving skills.

Table 4  
*Interventions for Child Witnesses of Intimate Partner Violence*

| Citation                                        | Sample and N                                              | Treatment (n)                                                                | Outcome measures                            | Posttreatment follow-up | Outcome                                                                                                                                                                                                                               | Treatment dropout                          | Follow-up attrition                |
|-------------------------------------------------|-----------------------------------------------------------|------------------------------------------------------------------------------|---------------------------------------------|-------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------|------------------------------------|
| Lieberman, Van Horn, Ippen, & Ghosh Ippen, 2005 | 75 preschool child-mother dyads exposed to IPV            | Child-parent psychotherapy (42) vs. community treatment-case management (33) | CBCL, SCL-90, CAPS                          | Posttreatment only      | Significant Group $\times$ Time interactions with children's behavior problems, PTSD symptoms, diagnostic status, and mother's avoidance; trend toward decreasing mother's PTSD and general distress at end of 1 year of intervention | Not reported                               | 14% tx vs. 12% comparison          |
| Lieberman, Ghosh Ippen, & Van Horn, 2006        | 75 preschool child-mother dyads exposed to IPV            | Child-parent psychotherapy (42) vs. community treatment-case management (33) | CBCL, SCL-90, CAPS                          | 6 months                | Improvement in children's behavior problems and maternal symptoms                                                                                                                                                                     | Not reported                               | 33% not available for follow-up    |
| Graham-Berman et al., 2007                      | 181 children 6–12 years old exposed to IPV                | CO (62) vs. CM (61) vs. wait-listed comparison (58)                          | CBCL, Attitudes About Family Violence Scale | 8 months                | CM condition had greatest improvement in externalizing scores, but CO improved compared with wait-listed group at posttest                                                                                                            | 18%, not randomized                        | 6% intervention vs. 18% comparison |
| Cohen et al., 2004                              | 237 sexually abused children: 90% PTSD, 58% IPV witnesses | TF-CBT (114) vs. CCT (115)                                                   | K-SADS-PL-PTSD, CBCL, CDI, CAPS, CSBI       | Posttreatment only      | TF-CBT associated with decrease in PTSD symptoms (19% posttest); CCT (46% posttest)                                                                                                                                                   | 11% of those enrolled attended <2 sessions | 12% TF-CBT vs. 10% CCT             |

*Note.* IPV = interpersonal violence; PTSD = posttraumatic stress; tx = treatment; CBCL = Child Behavior Checklist; SCL-90 = Symptom Checklist; CAPS = Clinician-Administered PTSD Scale; TF-CBT = trauma-focused cognitive-behavioral therapy; CCT = child-centered therapy; CO = child-only treatment; CM = child-and-mother treatment; K-SADS-PL-PTSD = Kiddie Schedule for Affective Disorders and Schizophrenia—Present and Lifetime Version, PTSD section; CDI = Child Depression Inventory; CSBI = Child Sexual Behavior Inventory.

choeducation about family violence, surveyed children's attitudes about families, and addressed their social emotional adjustment. Mothers' groups focused on parenting competence and understanding the impact of violence on children. This was a community sample, with 17% of the mothers and children still living with the abusive partner. Children whose mothers were seen concurrently showed the greatest reduction in externalizing symptoms (Graham-Bermann, Lynch, Banyard, Devoe, & Halabu, 2007).

A third treatment, trauma-focused cognitive-behavioral therapy (TF-CBT), has been the most vigorously studied and widely disseminated. In a randomized controlled trial with sexually abused children, 58% of whom also had a history of witnessing domestic violence, TF-CBT was associated with significantly better outcomes than supportive child-centered therapy (Cohen, Deblinger, Mannarino, & Steer, 2004). TF-CBT comprises specific modules including psychoeducation; expressing feelings; recognizing the relationship among thoughts, feelings, and behaviors; learning relaxation skills; gradual exposure; cognitive processing of the abuse experience; joint parent-child sessions; and parent management training to address behavioral problems (Cohen et al., 2004). TF-CBT is designed to be provided in 12–18 sessions, and caregiver involvement is important for treatment success.

*Summary of child-witness to IPV treatments.* Several treatments have shown promising effectiveness data, with conjoint treatment of mother and child being the most effective. These treatments primarily have been implemented with families in which the mother and child were no longer living with the perpetrator, with maternal substance abuse also an exclusion criterion, limiting the generalizability of these treatments in “real-world” settings.

## Discussion

Overall, results of this treatment review indicate a lack of research evidence for the broad, long-term effectiveness of many of the most common treatments provided for victims and perpetrators of IPV, including the Duluth model for perpetrators and shelter-advocacy approaches for working with victims of domestic violence. According to partner reports, rates of recidivism in most perpetrator- and partner-focused treatments are approximately 20%–30% within 6 months, regardless of intervention strategy used. This rate is comparable to the rate reported in studies examining the efficacy of mandatory arrest in deterring subsequent family violence.

Much more attention needs to be paid to the question of, “Which treatment for whom?” Blanket policies requiring specific treatment approaches for all male batterers are not effective. Assessment of individual treatment needs would allow for a better fit between individual batterers and their court-mandated treatment. While not initially developed to target IPV, behavioral couples therapy (BCT)—which integrates substance abuse treatment approaches, couples therapy, and CBT coping skills—appears to be an effective strategy for IPV cases in which one or both partners have a comorbid substance use disorder. BCT had the lowest rates of recidivism (18%) and treatment dropout (14%) compared with the other treatments reviewed for batterers. Given the high comorbidity between IPV and substance abuse problems, further systematic evaluation of this sort of integrated treatment approach appears warranted.

Advocacy interventions for victims of IPV result in increased feelings of safety and support and some short-term reductions in violence. Manualized dyadic or concurrent child-parent trauma-focused interventions (e.g. CPP and TF-CBT) have been shown to reduce symptoms in both children and their caregivers. Thus, incorporation or coordination of advocacy for victims and dyadic parent-child trauma-focused treatment, along with batterer intervention, may yield the best overall outcomes for families impacted by IPV. Instead, families are often referred to a variety of providers in multiple settings. The courts may mandate attendance at a batterers' group, substance abuse treatment, and a parenting class for a perpetrator of IPV. Additionally, child protective services may request that the victim-mother attend her own individual treatment and a parenting class, as well as bring her children for their own individual treatments. Often, these services are provided by a variety of agencies in different locations and are not well coordinated. Better integration of treatment approaches in one location (e.g. substance abuse, batterer, and parenting treatment for perpetrators) is needed.

Dropout is a significant problem in most treatment studies for batterers. In substance abuse treatment studies, intervention completion has been increased by incorporating motivational enhancement therapy (MET) techniques (Carroll & Onken, 2005). The goal of MET is to resolve ambivalence concerning whether or not the client has a problem and to increase motivation to change. There are five main strategies to motivational interviewing (Irons, 2006): (a) express empathy, (b) develop discrepancies, (c) avoid argumentation, (d) roll with resistance, and (e) support self-efficacy. Examination of these approaches in IPV treatments is warranted, with the parenting role a potentially valuable focus of MET interventions, as research suggests most batterers report an attachment to their children and an awareness of the negative impact of their violence on them (Baker, Perrilla, & Norris, 2001; Israel & Stover, in press).

Most IPV victims stay with or return to the batterer (Lerner & Kennedy, 2000). If partners separate, visitation is an ongoing issue, with one study finding that preschool-aged children who had limited contact with their previously violent fathers had higher levels of internalizing symptoms than children who had frequent (at least weekly) contact, even after controlling for the severity of violence exposure (Stover, Van Horn, Turner, Cooper, & Lieberman, 2003). Focusing on the perpetrators' role as parents in therapy, in addition to enhancing motivation for treatment, may also help to improve child outcomes. This proposition is supported by the promising results of dyadic treatment approaches in working with mothers and children exposed to IPV (Lieberman et al., 2006), and the efficacy of parent-child interaction therapy when used with physically abusive parents and their children (Chaffin et al., 2004; Toth, Maughan, Manly, Spagnola, & Cicchetti, 2002).

## Study Limitations

Although every attempt was made to do a thorough review of all available studies, it is possible that computerized literature searches missed relevant research. Given the well-documented “file drawer phenomenon”—the failure to publish negative studies—it can be assumed that the published literature captures only a subset of all research conducted in this area. In addition, implementation of treatment with batterers requires the use of

forensic tools to determine risk, coordination of treatment with the legal system or the child protective services system, and careful incorporation of safety planning to assure the well-being of victims and children. Unfortunately, the breadth of material covered in this review did not permit discussion of these additional important topics relevant in implementing and investigating IPV interventions.

### Summary and Closing Remarks

Extant interventions have limited effect on repeat violence, with most treatments reporting minimal benefit above arrest alone. The results of this treatment review indicate a lack of research evidence for the effectiveness of many of the most common treatments provided for victims and perpetrators of IPV, including the Duluth model for perpetrators and shelter-advocacy approaches for victims. According to partner reports, rates of recidivism in most perpetrator- and partner-focused treatments are approximately 30% within 6 months, regardless of intervention strategy used. Emerging data supports the integration of empirically validated substance abuse, couples, and trauma-focused interventions into IPV treatments. However, considerably more work is needed in this area.

### References

- Adams, D. A. (1988). Counseling men who batter: Pro-feminist analysis of five treatment models. In M. Bograd & K. Yllo (Eds.), *Feminist perspectives on wife abuse* (pp. 176–199). Newbury Park: Sage.
- Babcock, J. C., Green, C. E., & Robie, C. (2004). Does batterers' treatment work? A meta-analytic review of domestic violence treatment. *Clinical Psychology Review*, 23, 1023–1053.
- Baker, C. K., Perilla, J. L., & Norris, F. H. (2001). Parenting stress and parenting competence among Latino men who batter. *Journal of Interpersonal Violence*, 16, 1139–1157.
- Brannen, S. J., & Rubin, A. (1996). Comparing the effectiveness of gender-specific and couples groups in a court-mandated spouse abuse treatment program. *Research on Social Work Practice*, 6, 405–424.
- Brookoff, D., O'Brien, K. K., Cook, C. S., Thompson, T. D., & Williams, C. (1997). Characteristics of participants in domestic violence. Assessment at the scene of domestic assault. *Journal of the American Medical Association*, 277, 1369–1373.
- Bybee, D., & Sullivan, C. M. (2005). Predicting re-victimization of battered women 3 years after exiting a shelter program. *American Journal of Community Psychology*, 36, 85–96.
- Campbell, J., Jones, A. S., Dienemann, J., Kub, J., Schollenberger, J., O'Campo, P., et al. (2002). Intimate partner violence and physical health consequences. *Archives of Internal Medicine*, 162, 1157–1163.
- Carroll, K. M., & Onken, L. S. (2005). Behavioral therapies for drug abuse. *American Journal of Psychiatry*, 162, 1452–1460.
- Carter, L. S., Weithorn, L. A., & Behrman, R. E. (1999). Domestic violence and children: Analysis and recommendations. *Future of Children*, 9, 4–20.
- Chaffin, M., Silovsky, J. F., Funderburk, B., Valle, L. A., Brestan, E. V., Balachova, T., et al. (2004). Parent-child interaction therapy with physically abusive parents: Efficacy for reducing future abuse reports. *Journal of Consulting and Clinical Psychology*, 72, 500–510.
- Cohen, J. A., Deblinger, E., Mannarino, A. P., & Steer, R. A. (2004). A multisite, randomized controlled trial for children with sexual abuse-related PTSD symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry*, 43, 393–402.
- Cohen, J. A., Mannarino, A. P., & Deblinger, E. (2006). *Treating trauma and traumatic grief in children and adolescents* (1st ed.). Guilford, CT: The Guilford Press.
- Davis, R. C., Maxwell, C. D., & Taylor, B. G. (2006). Preventing repeat incidents of family violence: Analysis of data from three field experiments. *Journal of Experimental Criminology*, 2, 183–210.
- Dunford, F. W. (2000). The San Diego Navy experiment: An assessment of interventions for men who assault their wives. *Journal of Consulting and Clinical Psychology*, 68, 468–476.
- Dutton, M. A., Green, B. L., Kaltman, S. I., Roesch, D. M., Zeffiro, T. A., & Krause, E. D. (2006). Intimate partner violence, PTSD, and adverse health outcomes. *Journal of Interpersonal Violence*, 21, 955–968.
- Easton, C. J., Swan, S., & Sinha, R. (2000). Prevalence of family violence in clients entering substance abuse treatment. *Journal of Substance Abuse Treatment*, 18, 23–28.
- Eisenstat, S. A., & Bancroft, L. (1999). Domestic violence. *New England Journal of Medicine*, 341, 886–892.
- Fals-Stewart, W., Kashdan, T. B., O'Farrell, T. J., & Birchler, G. R. (2002). Behavioral couples therapy for drug-abusing patients: Effects on partner violence. *Journal of Substance Abuse Treatment*, 22, 87–96.
- Feder, L., & Dugan, L. (2002). A test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward experiment. *Justice Quarterly*, 19, 343–375.
- Feder, L., & Wilson, D. B. (2005). A meta-analytic review of court-mandated batterer intervention programs: Can courts affect abusers' behavior? *Journal of Experimental Criminology*, 1, 239–262.
- Ford, D. A., & Regoli, M. J. (1993). The criminal prosecution of wife batterers: Process, problems, and effects. In N. Z. Hilton (Ed.), *Legal responses to wife assault* (pp. 127–164). Newbury Park, CA: Sage.
- Golding, J. M. (1999). Intimate partner violence as a risk factor for mental disorders: A meta-analysis. *Journal of Family Violence*, 14, 99–132.
- Gortner, E. T., Gollan, J. K., & Jacobson, N. S. (1997). Psychological aspects of perpetrators of domestic violence and their relationships with the victims. *Psychiatric Clinics of North America*, 20, 337–351.
- Graham-Bermann, S. A., Lynch, S., Banyard, V., Devoe, E. R., & Halabu, H. (2007). Community-based intervention for children exposed to intimate partner violence: An efficacy trial. *Journal of Consulting and Clinical Psychology*, 75, 199–209.
- Harris, R., Savage, S., Jones, T., & Brooke, W. (1988). A comparison of treatments for abusive men and their partners within a family-service agency. *Canadian Journal of Community Mental Health*, 7, 147–155.
- Irons, B. L. (2006). Alcohol use disorders: A clinical update. *Adolescent Medicine Clinics*, 17, 259–282.
- Israel, E., & Stover, C. S. (in press). Intimate partner violence: The impact of the relationship between perpetrators and children who witness violence. *Journal of Interpersonal Violence*.
- Kendall-Tackett, K. A. (Ed.) (2004). *Health consequences of abuse in the family: A clinical guide for evidenced-based practice*. Washington, DC: American Psychological Association.
- Lerner, C. F., & Kennedy, L. T. (2000). Stay-leave decision making in battered women: Trauma, coping, and self-efficacy. *Cognitive Therapy and Research*, 24, 215–232.
- Lieberman, A. F., Ghosh Ippen, C., & Van Horn, P. (2006). Child-parent psychotherapy: 6-month follow-up of a randomized controlled trial. *Journal of American Academy of Child and Adolescent Psychiatry*, 4, 913–918.
- Lieberman, A. F., & Van Horn, P. (2004). *Don't hit my mommy: A manual for child-parent psychotherapy with young witnesses of family violence*. Washington, DC: Zero to Three.
- Lieberman, A. F., Van Horn, P., & Ghosh Ippen, C. (2005). Toward evidence-based treatment: Child-parent psychotherapy with preschoolers exposed to marital violence. *Journal of the American Academy of Child and Adolescent Psychiatry*, 44, 1241–1248.
- Mauricio, A. M., Tein, J. Y., & Lopez, F. G. (2007). Borderline and

- antisocial personality scores as mediators between attachment and intimate partner violence. *Violence and Victims*, 22, 139–157.
- Maxwell, C. D., Garner, J. H., & Fagan, J. A. (2001). *The effects of arrest on intimate partner violence: new evidence from the spouse assault replication program*. Retrieved June 16, 2008, from the National Criminal Justice Reference Service: <http://www.ncjrs.gov/txtfiles1/nij/188199.txt>
- McCloskey, L. A. P., & Walker, M. (2000). Posttraumatic stress in children exposed to family violence and single-event trauma. *Journal of American Academy of Child and Adolescent Psychiatry*, 39, 108–115.
- McFarlane, J., Soeken, K., & Wiist, W. (2000). An evaluation of interventions to decrease intimate partner violence to pregnant women. *Public Health Nursing*, 17, 443–451.
- O'Farrell, T. & Fals-Stewart, W. (2006). Behavioral couples treatment for alcoholism and drug abuse. New York: Guilford Press.
- O'Leary, K. D., Heyman, R. E., & Neidig, P. H. (1999). Treatment of wife abuse: A comparison of gender-specific and conjoint approaches. *Behavior Therapy*, 30, 475–505.
- Palmer, S. E., Brown, R. A., & Barrera, M. E. (1992). Group treatment program for abusive husbands: Long-term evaluation. *American Journal of Orthopsychiatry*, 62, 276–283.
- Pence, E. & Paymar, M. (1993). *Education groups for men who batter: The Duluth Model*. New York: Springer.
- Sherman, L. W., & Berk, R. A. (1984). The specific deterrent effects of arrest for domestic assault. *American Sociological Review*, 49, 261–272.
- Stover, C. S., Berkman, M., Desai, R., & Marans, S. (2008). The efficacy of a police–advocacy home-visit intervention: 12-month follow-up data. Manuscript submitted for publication.
- Stover, C. S., Van Horn, P., Turner, R., Cooper, B., & Lieberman, A. F. (2003). The effects of father visitation on preschool-aged witnesses of domestic violence. *Journal of Interpersonal Violence*, 18, 1149–1166.
- Sullivan, C. M., & Bybee, D. I. (1999). Reducing violence using community-based advocacy for women with abusive partners. *Journal of Consulting and Clinical Psychology*, 67, 43–53.
- Sullivan, C. M., Campbell, R., Angelique, H., Eby, K., & Davidson, W. (1994). An advocacy intervention program for women with abusive partners: Six-month follow-up. *American Journal of Community Psychology*, 22, 101–122.
- Taylor, B. G., Davis, R. C., & Maxwell, C. D. (2001). The effects of group batterer treatment: The Brooklyn experiment. *Justice Quarterly*, 18, 171–201.
- Tjaden, P., & Thoennes, N. (2000). *Full report of the prevalence, incidence, and consequences of violence against women: Research report* (NIJ Rept. No. NCJ 183781). Washington, DC: U.S. Department of Justice, National Institute of Justice.
- Toth, S. L., Maughan, A., Manly, J. T., Spagnola, M., & Cicchetti, D. (2002). The relative efficacy of two interventions in altering maltreated preschool children's representational models: Implications for attachment theory. *Development and Psychopathology*, 14, 877–908.
- Watts, C., & Zimmerman, C. (2002, April 6). Violence against women: Global scope and magnitude. *Lancet*, 359, 1232–1237.
- Wilt, S., & Olson, S. (1996). Prevalence of domestic violence in the United States. *Journal of the American Medical Women's Association*, 51, 77–82.

Received October 5, 2007

Revision received April 28, 2008

Accepted May 6, 2008 ■

### Clarification Notice

We wish to clarify the relationship between “Racial Microaggressions in the Life Experience of Black Americans” by Derald W. Sue, Christina M. Capodilupo, and Aisha M. B. Holder (*Professional Psychology: Research and Practice*, 2008, Vol. 39, No. 3, pp. 329–336) and “Racial Microaggressions Against Black Americans: Implications for Counseling” by Derald W. Sue, Kevin L. Nadal, et al. (*Journal of Counseling and Development*, 2008, Vol. 86, No. 3 pp. 330–338). These two articles are based on the same sample of subjects and set of interviews; however, separate qualitative analyses by different teams of researchers were performed on the transcripts of the interviews. The first study investigated racial microaggressive dynamics, processes, and their detrimental consequences for African Americans, whereas the second study explored the universe of hidden demeaning racial microaggressive themes. In the second article, which did not mention the sample overlap, a few descriptive sentences from the first article, primarily in the Method-section, were repeated verbatim and without citation from the earlier study. We apologize for these oversights.

Derald Wing Sue

DOI: 10.1037/a0016335



# Improving Batterer Intervention Programs Through Theory-Based Research

Gregory L. Stuart, PhD

Jeff R. Temple, PhD

Todd M. Moore, PhD

INTIMATE PARTNER VIOLENCE (IPV) IS A MAJOR PUBLIC HEALTH concern, with at least 1.3 million women abused annually in the United States.<sup>1</sup> In the 1980s, states began passing mandatory arrest laws that required police officers responding to domestic violence calls to make arrests when there was evidence of probable cause of violence perpetration. The proliferation of these laws inevitably led to increasing numbers of male batterers (defined here as men who are arrested for aggression against a partner) entering the criminal justice system. In a rush to address the needs of abused women, states began implementing batterer intervention programs attempting to reduce IPV recidivism. Most jurisdictions require some intervention postarrest for partner-violent men, and the majority of men in these programs have been court ordered to attend.<sup>2</sup> Unfortunately, the overwhelming demand for and expansion of these programs outpaced research efforts assessing their efficacy. An increasing proportion of women are being arrested for partner violence and mandated to attend batterer intervention programs. Although our focus is on male offenders, additional attention should be directed toward evaluating and improving these programs for women, particularly considering that they may have distinct needs from men and would likely benefit from a different form of intervention.

## Efficacy of Batterer Intervention Programs

Batterer intervention programs are generally conducted in a group format ranging from 12 to 52 weeks in duration.<sup>2,3</sup> They are typically considered a combination of rehabilitation and punishment, with an emphasis on the safety of the abused partner. Although a variety of approaches exist, program content is typically derived from feminist or cognitive behavioral therapy (CBT) models. Feminist-based programs (eg, the Duluth model<sup>4</sup>) target men's views of women and men's belief that they should be able to control their partners. Interventions are designed to help men examine their sexist assumptions and patriarchal beliefs about relationships. Men are asked to critically examine the various methods they use to control their partners as well as how society may sanction such actions to develop strategies to eliminate violent behavior. In contrast, CBT groups are based on assumptions that violent men have deficits in anger control and relationship and communication skills. The assumption is that violent men who en-

counter relationship conflicts may be unable to engage in conflict-reducing communication, resulting in increased anger and a greater risk for engaging in violence. In practice, it may be difficult to label a particular group as feminist or CBT, as batterer programs often blend these theoretical conceptualizations and intervention techniques.<sup>2,3,5</sup>

In most states, men arrested for violence against their intimate partners are mandated to attend a batterer intervention program, where they are educated on the misuse of power and control over partners and/or taught skills to communicate more effectively, manage anger, and solve relationship problems. However, there is scant empirical evidence that these programs reduce recidivism. Numerous studies, qualitative reviews, and meta-analyses have repeatedly arrived at a similar conclusion: batterer intervention programs have a small, often nonsignificant effect in reducing partner violence.

In a meta-analysis of 22 batterer intervention studies in which a comparison group was included (eg, program dropouts, nonequivalent controls), Babcock et al<sup>3</sup> concluded that recidivism was 5% less likely by men arrested and referred to a batterer intervention program than by men arrested and sanctioned without intervention. The effectiveness of these programs in reducing recidivism was consistently low regardless of the source of information on violence (eg, police arrest records vs partner reports) or the theory emphasized in the intervention (eg, feminist vs CBT).

Similarly, Feder and Wilson<sup>5</sup> conducted a meta-analysis using only the 10 most rigorous studies. In controlled studies that involved randomization of participants and official reports to measure recidivism, they found a 7% decrease in recidivism beyond traditional criminal justice interventions, such as probation or community service. When using partner reports as the outcome measure, which is arguably a higher and more accurate estimate of violence recidivism, they found no benefit of batterer intervention programs.

## Reasons for the Ineffectiveness of Batterer Intervention Programs

One potential reason for the ineffectiveness of batterer intervention programs is that partner-violent men are typically man-

**Author Affiliations:** The Warren Alpert Medical School of Brown University and Butler Hospital, Providence, Rhode Island (Drs Stuart and Temple); and Department of Psychology, University of Tennessee, Knoxville (Dr Moore). Dr Temple is now with the Department of Obstetrics and Gynecology, University of Texas Medical Branch, Galveston.

**Corresponding Author:** Gregory L. Stuart, PhD, The Warren Alpert Medical School of Brown University and Butler Hospital, Psychiatry and Human Behavior, 345 Blackstone Blvd, Providence, RI 02906 (gregory\_stuart@brown.edu).

dated to attend by the judicial system<sup>2</sup> and may be unwilling or unmotivated to accept responsibility for being violent. Men attending these programs often perceive having been “forced” into an unwarranted intervention because their partners either lied to the police about violence and/or the violence was bidirectional or partner initiated. Many batterers minimize the severity of their abuse and often deny it completely.<sup>4,6</sup> Engaging men in the intervention process under mandatory conditions while surrounded by other men who may resent and blame their partner, the system, or both for the current situation reduces the likelihood of significant change. Second, batterer intervention programs receive inadequate funding, resulting in limited resources and overworked clinicians who often have minimal training and lack advanced professional degrees.<sup>7,8</sup> Third, the intervention is seldom tailored to meet the clients’ specific needs, such as addressing substance misuse, other psychological problems, or both. Thus, the context, motivation, and overall individual variability are not considered when treating batterers. For example, a severely violent serial offender with extensive comorbid psychopathology may require a different approach than a one-time offender of low-level violence with no comorbidity. Fourth, due to the rapid expansion of batterer intervention programs and the rush to implement them by the criminal justice system, these programs were already in wide use before rigorously evaluating their efficacy.

### Suggestions to Increase Effectiveness

Before making radical programmatic changes, a systematic, theoretically based, and empirically driven approach is needed to improve the quality of batterer intervention programs. Although the efficacy has been negligible, these programs may result in some benefits over and above traditional criminal justice interventions. Babcock et al<sup>3</sup> estimated that a 5% decrease in recidivism translates into the elimination of violence for approximately 42 000 women. However, policy makers, researchers, and batterer program administrators should acknowledge the preponderance of evidence that these programs lack sufficient effectiveness, and that significant effort must be devoted to identifying methods to improve them.<sup>7-9</sup> To this end, the following suggestions may be considered as first steps.

**Motivational Strategies.** Murphy and Baxter<sup>6</sup> propose using the transtheoretical model of behavior change<sup>10</sup> as the backdrop to using motivational interviewing techniques to modify violent behavior. The transtheoretical model postulates that there are at least 5 stages in the process of change, ranging from precontemplation (person is not considering change) through contemplation, preparation, action, and maintenance (identifying methods to avoid relapse). Motivational interviewing strives to meet clients at their current state of readiness and uses nonconfrontational strategies to assist individuals in eliciting their own reasons for change.<sup>11</sup> It is assumed that developing a supportive working relationship with clients will reduce defensiveness and increase willingness to

explore the need for change over more confrontational approaches. This may be a viable addition to current batterer intervention programs that typically use confrontational techniques to modify men’s behavior. A series of studies found that most men entering batterer intervention programs are in the early stages of change<sup>12</sup> and that motivational interviewing increased session attendance and reduced posttreatment IPV.<sup>13</sup> This approach can be integrated into current batterer intervention programs relatively easily and may be particularly important given that these programs often assume that the clients entering the intervention are ready for change.<sup>12</sup>

**Tailored Treatment.** Some researchers have proposed that men engaging in IPV constitute a heterogeneous group and that there may be subtypes of partner-violent men, each with a different etiology of violence.<sup>7,14</sup> Thus, tailoring treatments to meet the needs of specific subtypes of violent men, or the needs of individual clients, might improve therapy efficacy. Although research on subtypes of violent perpetrators is in its infancy, preliminary research suggests that some subtypes may have better IPV outcomes than others.<sup>8,12</sup> At minimum, one way to improve batterer intervention program outcomes would be to conduct thorough assessments of the particular needs of each individual to determine the optimal set of interventions to maximize outcomes.<sup>7</sup> Although batterer programs may not have the resources to fully incorporate these additional interventions, ideally they would conduct multiple assessments of their clients’ needs throughout the course of the intervention, refer to collateral care as appropriate, and encourage participants to follow through with treatment recommendations.

**Substance Abuse Treatment.** There is an abundance of theoretical and empirical evidence linking substance use to IPV. Addictive behaviors are overrepresented in populations of batterers. In 1 study,<sup>15</sup> 68% of men attending batterer intervention programs met criteria for hazardous drinking, 53% had a probable alcohol diagnosis, and 42% were likely alcohol dependent. Moreover, 54% of the men attending these programs had used an illicit substance in the past year<sup>16,17</sup> and 31% of the sample had a probable drug use disorder.<sup>15</sup> There appears to be a temporal association between substance use and violence. For example, research has demonstrated that men attending batterer programs were 20 times more likely to assault their partners on a heavy drinking day relative to a nondrinking day.<sup>18</sup> Despite this evidence and the finding that men in batterer intervention programs with addictions have significantly worse outcomes than those without addictions,<sup>19</sup> research has shown that only 3% of men arrested for IPV were court mandated to also attend substance abuse treatment.<sup>20</sup>

How can batterer intervention programs be modified to address the special needs of the large proportion of male batterers who have addictions? Although addictions do not justify or excuse violence, it is unlikely that an individual with an active substance use disorder will effectively learn and practice the skills taught in a batterer intervention program. Substance use may precipitate relationship conflict

and may compromise an individual's ability to reflect upon and use the skills taught in these programs, which may increase the risk of aggression. Although the finding did not reach statistical significance, 1 study reported that the probability of violence recidivism subsequent to batterer intervention was reduced by 30% to 40% if the individual obtained substance abuse treatment.<sup>19</sup> In addition, substance abuse interventions are associated with substantial reductions in IPV, even when the substance use interventions did not specifically target relationship aggression.<sup>21,22</sup> Moreover, relapsed alcoholics perpetrate more IPV than remitted alcoholics after treatment, with remitted alcoholics showing IPV perpetration at levels close to the general population.<sup>21</sup> Given this evidence and the high comorbidity of substance use disorders in batterers, it is likely that many men arrested for violence would benefit from integrating addictions treatments into batterer intervention programs or obtaining concurrent substance abuse treatment.

**Conjoint (Couples) Treatment.** Although justifiably controversial, couple-level interventions have demonstrated some efficacy in reducing IPV.<sup>23,24</sup> Although couples approaches for men arrested for IPV are prohibited in many states,<sup>2,3,5,7</sup> previous violence intervention studies that used couples did not find any additional danger to the women involved.<sup>23</sup> Nonetheless, several caveats need to be considered before implementing conjoint therapy. This method may not be appropriate unless there is a history of low or moderate levels of violence, the woman independently agrees to participate and does not express fear of negative consequences for discussing the relationship and violence, and both partners commit to avoiding additional physical aggression.<sup>24</sup> Under these conditions and if the couple plans to stay together, conjoint treatment offers several advantages over traditional batterer intervention techniques. IPV occurs within a dyad and is often predicted by relationship discord, and therefore treating the male in isolation may have less long-term impact if relationship issues are not also addressed.

Another argument in favor of conjoint therapy is the fact that a majority of violent relationships are characterized by mutual violence in which both partners are aggressive toward the other.<sup>25</sup> Although women are far more likely than men to be the recipients of severe and injurious forms of violence and to report poor mental health as a consequence of IPV, violence often begets violence, so the cessation of aggression within a relationship may be less likely unless accomplished at the couple level.<sup>23</sup> For carefully selected clients, couples approaches may be helpful adjuncts to batterer intervention programs, may be beneficial subsequent to traditional batterer intervention, or in rare cases may be useful in lieu of batterer intervention.

## Conclusions

Each of the suggestions to increase the effectiveness of batterer intervention programs are empirical questions worth

investigating. The recommendations should be examined with methodologically rigorous research designs to establish the efficacy of the program modifications before widespread dissemination.

**Financial Disclosures:** None reported.

**Funding/Support:** This work was supported in part by grants R01AA014193 and R01AA016315 from the National Institute on Alcohol Abuse and Alcoholism awarded to Dr Stuart.

**Role of the Sponsor:** The National Institute on Alcohol Abuse and Alcoholism had no role in the preparation, review, or approval of the manuscript.

## REFERENCES

1. National Center for Injury Prevention and Control. *Costs of Intimate Partner Violence Against Women in the United States*. Atlanta, GA: Centers for Disease Control and Prevention; 2003.
2. Healey K, Smith C, O'Sullivan C. *Batterer Intervention: Program Approaches and Criminal Justice Strategies*. Washington, DC: US Dept of Justice; 1998.
3. Babcock JC, Green CE, Robie C. Does batterers' treatment work? a meta-analytic review of domestic violence treatment. *Clin Psychol Rev*. 2004;23(8):1023-1053.
4. Pence E, Paymar M. *Education Groups for Men Who Batter: The Duluth Model*. New York, NY: Springer; 1993.
5. Feder L, Wilson D. A meta-analytic review of court-mandated batterer intervention programs. *J Exp Criminol*. 2005;1(2):239-262.
6. Murphy CM, Baxter VA. Motivating batterers to change in the treatment context. *J Interpers Violence*. 1997;12(4):607-619.
7. Stuart RB. Treatment for partner abuse: time for a paradigm shift. *Prof Psychol Res Pract*. 2005;36(3):254-263.
8. Holtzworth-Munroe A. Standards for batterer treatment programs: how can research inform our decisions? *J Aggress Maltreat Trauma*. 2001;5(2):165-180.
9. Stuart GL. Improving violence intervention outcomes by integrating alcohol treatment. *J Interpers Violence*. 2005;20(4):388-393.
10. Prochaska JO, DiClemente CC. *The Transtheoretical Approach: Crossing the Traditional Boundaries of Therapy*. Homewood, IL: Dow Jones Irwin; 1984.
11. Miller WR, Rollnick S. *Motivational Interviewing: Preparing People for Change*. 2nd ed. New York, NY: Guilford Press; 2002.
12. Eckhardt C, Holtzworth-Munroe A, Norlander B, Sibley A, Cahill M. Readiness to change, partner violence subtypes, and treatment outcomes among men in treatment for partner assault. *Violence Vict*. In press.
13. Taft CT, Murphy CM, Musser PH, Remington NA. Personality, interpersonal, and motivational predictors of the working alliance in group cognitive-behavioral therapy for partner violent men. *J Consult Clin Psychol*. 2004;72(2):349-354.
14. Holtzworth-Munroe A, Stuart GL. Typologies of male batterers: three subtypes and the differences among them. *Psychol Bull*. 1994;116(3):476-497.
15. Stuart GL, Moore TM, Kahler CW, Ramsey SE. Substance abuse and relationship violence among men court-referred to batterer intervention programs. *Subst Abuse*. 2003;24(2):107-122.
16. Moore TM, Stuart GL. Illicit substance use and intimate partner violence among men in batterers' intervention. *Psychol Addict Behav*. 2004;18(4):385-389.
17. Stuart GL, Meehan J, Temple JR, et al. The role of drug use in a conceptual model of intimate partner violence in men and women arrested for domestic violence. *Psychol Addict Behav*. In press.
18. Fals-Stewart W. The occurrence of partner physical aggression on days of alcohol consumption. *J Consult Clin Psychol*. 2003;71(1):41-52.
19. Jones AS, Gondolf EW. Time-varying risk factors for reassault among batterer program participants. *J Fam Violence*. 2001;16(4):345-359.
20. Goldkamp JS, Weiland D, Collins M, White M. *The Role of Drug and Alcohol Abuse in Domestic Violence and Its Treatment: Dade County's Domestic Violence Court Experience*. From the Executive Summary of a Crime and Justice Research Institute study funded by the National Institute of Justice; 1996.
21. O'Farrell TJ, Fals-Stewart W, Murphy M, Murphy CM. Partner violence before and after individually based alcoholism treatment for male alcoholic patients. *J Consult Clin Psychol*. 2003;71(1):92-102.
22. Stuart GL, Ramsey SE, Moore TM, et al. Reductions in marital violence following treatment for alcohol dependence. *J Interpers Violence*. 2003;18(10):1113-1131.
23. Stith SM, Rosen KH, McCollum EE. Treating intimate partner violence within intact couple relationships: outcomes of multi-couple versus individual couple therapy. *J Marital Fam Ther*. 2004;30(3):305-318.
24. O'Farrell TJ, Fals-Stewart W. *Behavioral Couples Therapy for Alcoholism and Drug Abuse*. New York, NY: Guilford; 2006.
25. Archer J. Sex differences in aggression between heterosexual partners: a meta-analytic review. *Psychol Bull*. 2000;126(5):651-680.



## THE EFFECTS OF A GROUP BATTERER TREATMENT PROGRAM: A RANDOMIZED EXPERIMENT IN BROOKLYN\*

**BRUCE G. TAYLOR\*\***  
National Institute of Justice

**ROBERT C. DAVIS\*\*\***  
Vera Institute of Justice

**CHRISTOPHER D. MAXWELL\*\*\*\***  
Michigan State University

---

\* This study was funded by a National Institute of Justice Grant (96-IJ-CX-0047) to Victim Services, New York, New York. This project was conducted while two of the authors (Bruce G. Taylor and Robert C. Davis) were employed with Victim Services, New York, New York. The views expressed are our own and do not necessarily represent the official positions of Victim Services, the National Institute of Justice, the Vera Institute of Justice or Michigan State University. Direct all correspondence to Dr. Bruce Taylor, National Institute of Justice, 810 Seventh Street, NW Washington, D.C. 20531-0001, [taylorb@ojp.usdoj.gov](mailto:taylorb@ojp.usdoj.gov).

\*\* Bruce Taylor is the deputy director of the Arrestee Drug Abuse Monitoring (ADAM) Program in the Office and Research and Evaluation at the National Institute of Justice (NIJ). His current intramural research projects at NIJ include evaluating a jail-based batterer treatment program through an experimental design; pilot testing a new domestic violence addendum for the ADAM Program; a multicity study examining HIV testing, risk, and prevention behaviors of arrestees as they relate to HIV and other sexually transmitted diseases; studies comparing the prevalence of illicit drug use among arrestees in seven countries; and studies examining different drug-testing kits and the validity of self-reports of drugs and alcohol. His most recent published research appears in *Women and Criminal Justice*, *Criminology*, *Violence and Victims*, *Journal of Personality and Social Psychology*, and *Journal of Drug Issues*. He received his doctorate from Rutgers University in 1996.

\*\*\* Robert Davis is a senior research associate at the Vera Institute of Justice, and a consultant for the American Bar Association. He is co-author of a recent book on crime prevention and principal editor of a recent book on crime victims. Currently he is working on a project funded by the Ford Foundation to promote democratic policing in emerging democracies. He is also conducting NIJ projects on innovative approaches to domestic violence prosecution and a national evaluation of victim services.

\*\*\*\* Christopher D. Maxwell is an assistant professor in the School of Criminal Justice at Michigan State University. His research interests include the social control and criminal justice processing of intimate violence, the efficacy of aggression and delinquency prevention programs, and the impact of social and ecological contexts on patterns of delinquency, crime, and criminal justice decision making. Current projects include several experiments testing criminal justice-based treatment programs for spouse abusers; a reanalysis of six experiments, each of which tested for the deterrent effect of arrest on spouse abusers; an assessment of how police handle domestic conflicts in the context of community policing; an evaluation of a community collaborative approach for response to domestic violence; and a study of the effects of various ecological contexts on individual-level judicial decision making. Dr. Maxwell is a graduate of Rutgers University (PhD, 1998; MA, 1994) and Indiana University-Bloomington (BA, 1990).

JUSTICE QUARTERLY, Vol. 18 No. 1, March 2001  
© 2001 Academy of Criminal Justice Sciences

Despite numerous evaluations of batterer treatment programs, most lack sufficient methodological rigor to yield valid answers about the programs' effectiveness. This paper presents results from an experimental evaluation in which 376 adult males convicted of domestic violence were randomly assigned to either a 40-hour batterer treatment program or 40 hours of community service that did not include any therapeutic treatment. We examined both official records and victims' reports of recidivism. Those assigned to the treatment program showed significantly lower recidivism, on the basis of all outcome measures from official records. Although victims' reports also recorded fewer failures among the batterers assigned to the treatment group, the differences in failure rates were not large enough to be statistically significant. Overall results suggest that therapeutic treatment for batterers may reduce domestic violence among convicted batterers who agree to this sentence.

Therapeutic treatment programs for batterers became a popular court sanction beginning in the early 1980s. As state and national policies increasingly promoted arrest and prosecution to control domestic violence (Buzawa and Buzawa 1996), this sanctioning method has become more important. Because of these changes, criminal courts have sanctioned an expanding pool of batterers, and judges have relied increasingly on group treatment programs as their sanction of choice despite victims' desires or degree of willingness to cooperate (Hanna 1996; Rebovich 1996). By the late 1990s nearly every state used batterer treatment programs; administrators estimated that nearly 80 percent of their clients were referred by the courts (Healey, Smith, and O'Sullivan 1998).

Because of the substantial growth in the number of batterer treatment programs, it is important to understand how effective these programs are in changing batterers' behavior. Although the number of female victims of intimate violence declined in the 1990s, an estimated 840,000 women annually were assaulted by their current or former spouses or boyfriends (Greenfeld et al. 1998). Furthermore, by 1996 nearly 25 million women — one-fourth of the adult female population — reported that they had experienced some type of physical victimization during their lifetime by a family member or intimate partner (Tjaden and Thoennes 1998). Therefore an intervention that also reduces the likelihood of future domestic violence will benefit many women over their lifetime. In addition, because many victims stay with their partners even after the batterer's arrest and conviction, it is essential to use effective programs that can change abusive behavior rather than simply delaying it during a period of incarceration. Also, some observers have argued that batterer treatment groups have the potential to create a "ripple effect" throughout the criminal justice system (Dutton 1986). For example, if such groups are effective, police may be more willing to make arrests and take domestic violence seriously,

prosecutors may be more willing to proceed with domestic violence cases, and judges may be less willing to use incarceration as the only sanction.

More than three dozen evaluations of batterer treatment programs have been conducted (for a recent review, see Davis and Taylor 1999). Although these evaluations show an evolution toward more rigorous science since the first studies in the early 1980s, many still lack sufficient methodological rigor to yield valid answers about the effectiveness of the programs. Our study is one of the most recent attempts to test batterer treatment using an experimental design that randomly assigns court-mandated batterers to treatment or to a control condition. In this study we address some methodological problems found in prior research, such as disentangling the effects of treatment from sample selection effects.

### **PRIOR RESEARCH ON BATTERER TREATMENT**

There is no lack of empirical studies on batterer treatment programs, as shown by at least six published reviews of more than three dozen published single-site evaluations (e.g., Eisikovits and Edleson 1989; Gondolf 1995; Rosenfield 1992; Saunders 1996; Tolman and Bennett 1990) and eight research reviews (Crowell and Burgess 1996; Davis and Taylor 1999; Dobash et al. 1995; Dutton 1988; Hamberger and Hastings 1993; Rosenbaum and O'Leary 1986; Saunders and Azar 1989; Tolman and Edelson 1995). This situation is deceptive, however, because only a handful of investigations can make any valid claim about differences between treated and untreated batterers. In the following section we review three generations of batterer treatment studies: Those which failed to use control groups and examined only batterers assigned to treatment programs, quasi-experiments, and those studies which assigned batterers randomly to treatment.

#### *Studies Without Control Groups*

The oldest and largest portion of empirical literature consists of studies that examine outcomes only among batterers assigned to treatment programs. This set includes studies that assess violence or other outcomes among participants in treatment programs only after treatment (single-group posttest-only designs), studies that measure violence among treatment participants both before and after treatment (single-group pretest/posttest designs), and studies that compare the violence of batterers who completed treatment with that of batterers who did not complete treatment, but were assigned to receive it. This literature contains more than two dozen examples. Studies of this type were important in developing this

area of research, and provided a foundation for the newer, stronger designs.

It is difficult, however, to interpret results from these nonexperimental studies. Single-group, posttest-only designs provide no reference point by which to judge whether treatment programs reduce violence. Single-group pretest/posttest designs that show a reduction in violence after batterers participate in a treatment program are problematic because research repeatedly has shown that domestic violence declines after the police are called, even if nothing else is done. In fact, research suggests that only about one-third of batterers commit another act of domestic violence within six months after the police intervene (e.g., Davis and Taylor 1997; Fagan et al. 1984; C.D. Maxwell 1998). In addition, among studies that compared outcomes between batterers who completed treatment and those who did not, the treated and the untreated (dropout) groups were not comparable before treatment. Palmer, Brown, and Barrera (1992) suggest that better attendance indicates a greater motivation to change, even before treatment.

### *Quasi-Experiments*

The next generation of evaluation studies includes quasi-experimental designs using nonequivalent matched groups. In at least four studies, batterers whom the court mandated to treatment were compared with batterers who received other interventions (Chen et al. 1989; Dobash et al. 1996; Dutton 1986; Harrell 1991). These studies usually are more rigorous than those without control groups because they examine larger samples, do not rely only on batterers' self-reports to determine new violence, and include follow-up periods of at least one-year.

Among the four quasi-experimental studies, three reported lower violence among treated than among untreated batterers (Chen et al. 1989; Dobash et al. 1996; Dutton 1986). The effect sizes, however, were not always statistically significant. Also, they depended on which outcome measures were examined and whether comparisons involved all men assigned to treatment or only those who completed a requisite number of sessions. In contrast to these three studies, Harrell (1991) found that men who completed treatment were more abusive after treatment than men in a control sample.

Although quasi-experiments are a step forward from studies without control groups, Palmer et al. (1992) point out that they are not reliable, unbiased estimates of treatment effects because we cannot know whether batterers assigned to treatment and those in control groups were equivalent before the treatment. In some

quasi-experiments (such as Dutton 1986 or Harrell 1991), we know that the control group comprised batterers more prone to recidivate than those in the treated group; thus the study favored the finding of treatment effects. Researchers may try to control statistically for any pretreatment differences between groups; these controls, however, may not include key factors related to outcomes.

### *Randomized Experiments*

The safest way to ensure that sample mean estimates are unbiased is by assigning batterers randomly to treatment and control groups. In 1992 Palmer and colleagues conducted the first batterers' treatment experiment, randomly assigning batterers to a true no-treatment control group. The number of subjects, however, was far smaller than was needed to detect treatment effects: 59 probationers were assigned by a "block random" procedure to either a 10-session psychoeducational group (combining group discussion with information) or a no-treatment control group. To measure outcomes, Palmer and her colleagues examined police reports six months after treatment; they found that recidivism rates (domestic physical abuse or serious threats) for the treatment group were just one-third those of the control group. Even with the small *N*, this difference was statistically significant. Although Palmer et al. (1992) attempted to measure violence on the basis of surveys of victims and batterers, low response rates and a small sample size precluded any analysis of recidivism based on interview data.

Dunford (1999) completed another randomized study. He assigned 861 domestic violence perpetrators who were in the Navy to one of four programs: group treatment for batterers, couples counseling, rigorous monitoring (periodic calls to victims and record checks), or safety planning for the victim (intended to approximate a no-treatment control). The results have not yet been published, however. Florida Atlantic University is conducting an experiment, but results are not yet available. In this study, the Broward County (Florida) domestic court judges randomly assigned more than 400 convicted male misdemeanor domestic violence offenders to either one-year probation and Duluth-like batterer treatment or a control group that received only one year of probation (Feder 1999).

### **PURPOSES OF THE PRESENT STUDY**

We agree with Fagan (1996) that randomized experiments raise fewer questions about internal validity than other research designs. Our study adds to the literature on randomized experiments in several ways. First, our planned sample size was based on

an examination of effect sizes described in the literature. We designed the study to test the treatment's effects on several indicators of violence and attitudes, including victims' reports, which were not used by Palmer et al. (1992) because of the small number of surveys completed by victims.

Second, whereas the Palmer et al. (1992) experiment included all batterers sentenced to probation, regardless of the batterers' willingness to enter treatment, our study involved only cases in which prosecutor, defendant, and judge agreed that treatment was appropriate.<sup>1</sup> Although our results may not be as generalizable as Palmer's, our study did not include batterers who were completely unmotivated.<sup>2</sup> This is a key point because it has been argued (e.g., Rosenfeld 1992) that treatment cannot be expected to work for individuals treated against their will. Essentially the Palmer et al. (1992) study tested cases more typically referred to court-mandated batterer groups. Our study tests a subset of court-ordered batterer treatment cases: that is, men who agreed to a sentence including batterer treatment.

Finally, our experiment included both six- and 12-month follow-up interviews with the victims. The six-month interviews were important because any treatment effect may last only while a batterer is in treatment. The 12-month follow-up interviews were held to determine whether any initial effect of treatment would lessen after batterers were no longer in treatment but still were under court control.

## METHOD

The study used an experimental design: 376 male criminal court defendants charged with assaulting their intimate female partners were assigned randomly to a 40-hour batterer treatment program or to 40 hours of community service without batterer education or treatment groups. In the community service program, participants were required to help clean local parks and public buildings.

---

<sup>1</sup> The victims in these cases were not consulted as to whether they felt treatment was appropriate. In our study, however, the victim agreed in all cases to cooperate with the prosecutor and provide testimony. If they were not willing to testify against the defendants, the prosecutor generally dropped their cases and they did not reach our sample.

<sup>2</sup> The men in our sample were not pure volunteers; they were offered a plea agreement that included batterer treatment. These men could have accepted or rejected the offer, but they would have gone to trial if they had rejected it. If found guilty, they would have faced a more severe sentence; if found not guilty, they might have received no sanction at all. Our recruitment process probably filtered out many batterers who would have been unmotivated treatment participants. Accordingly, we mean by the term *motivated* that the men in our sample were sufficiently motivated to choose treatment over an alternative, possibly more severe sanction.

Assignments were made during sentencing, after all parties (judge, prosecutor, and defendant) agreed to batterer treatment if it was available through the random assignment process. To provide several assessments, we interviewed the batterers and their victims about new violence on three occasions: at the time of sentencing, six months after sentencing, and 12 months after sentencing. In addition, 12 months after sentencing we searched police records for new complaints to the police by the victims and for arrests of the batterers.

### *Sample Characteristics*

The sample consisted of males convicted of a misdemeanor domestic violence offense (although some originally were charged with felonious assault) in Kings County (Brooklyn, NY), in which the judge, the prosecutor, and the defendant agreed in principle to accept batterer's treatment if the Alternatives to Violence (ATV) program accepted the defendant. This sample represented a small percentage of about 11,000 domestic violence defendants adjudicated (i.e., with dispositions other than dismissal) during intake between February 19, 1995 and March 1, 1996. We selected a total of 376 cases for the experiment, about 1.5 cases per day.

Nearly two-thirds (64 percent) of the selected defendants were charged with third-degree assault, a class A misdemeanor. Another 19 percent initially were charged with felonious assault, although these individuals later pled to a misdemeanor charge. The remaining 17 percent were charged with violating restraining orders, menacing, harassment, and other offenses. Conditional discharges were the most common disposition (76 percent of the individuals in the sample).<sup>3</sup> Twenty-two percent of the cases were adjourned in contemplation of dismissal, a form of pretrial diversion in which defendants' cases are dismissed and their records are expunged if they avoid arrest and adhere to judicial conditions for six months. Two percent of the sample received probation. Conditional discharges and probation place defendants under court control for one year, compared with a period of six months for most adjournments in contemplation of dismissal. We also investigated whether the treatment and the control groups received the same quality and quantity of court supervision and found no difference in the supervision received by the two groups.<sup>4</sup>

<sup>3</sup> Overall, in misdemeanor domestic violence cases in Kings County Criminal Court, about 25 percent are dismissed, 45 percent receive an adjournment in contemplation of dismissal, 25 percent receive a conditional discharge, and about 5 percent receive probation or a jail sentence.

<sup>4</sup> We found no statistically significant difference between the groups in the percentage of cases receiving different types of supervision. Among those assigned

All of the batterers (see Table 1) were males with a mean age of 33 years. Thirty-six percent were African-American, 28 percent Hispanic, and 21 percent West Indian. The remaining 16 percent were either white, Asian, or other. About one-third reported not having a high school diploma, another one-third reported earning a high-school diploma or GED, and the other one-third reported some formal education beyond high school. Approximately two-thirds (64 percent) of the men reported employment (either part-time or full-time) at the time of their arrest; just 40 percent said they had been employed continuously during the past year. The mean household income was approximately \$16,000 per year, 41 percent personally earned less than \$7,000, and 10 percent earned more than \$27,000.

**Table 1. Batterers' Characteristics, by Assigned Treatment**

|                               | Total    | ATV      | Control  |
|-------------------------------|----------|----------|----------|
| Total N of Assigned Batterers | 376      | 190      | 186      |
| Age (mean)                    | 33.0     | 32.7     | 33.3     |
| Race/Ethnicity                |          |          |          |
| African-American              | 36%      | 32%      | 40%      |
| Hispanic                      | 28%      | 30%      | 26%      |
| Western Carribean             | 21%      | 19%      | 22%      |
| White, Asian, or other        | 16%      | 19%      | 12%      |
| Education                     |          |          |          |
| No diploma/GED                | 38%      | 36%      | 40%      |
| High school diploma/GED       | 31%      | 31%      | 31%      |
| More than high school         | 31%      | 33%      | 30%      |
| Employed at Time of Arrest    | 64%      | 64%      | 63%      |
| Employed Throughout Year      | 40%      | 45%      | 34%      |
| Household Income (Mean)       | \$16,300 | \$16,500 | \$16,100 |
| Personal Income               |          |          |          |
| No income                     | 19%      | 20%      | 19%      |
| \$1.00 to \$7,500             | 25%      | 28%      | 22%      |
| \$7,501 to \$17,000           | 28%      | 28%      | 28%      |
| \$17,001 to \$27,000          | 17%      | 16%      | 19%      |
| Over \$27,000                 | 10%      | 9%       | 11%      |
| Relationship Status           |          |          |          |
| Married                       | 43%      | 44%      | 41%      |
| Separated/divorced            | 8%       | 8%       | 7%       |
| Living together               | 19%      | 18%      | 20%      |
| Intimate                      | 9%       | 9%       | 9%       |
| Past intimate                 | 25%      | 26%      | 25%      |
| Other                         | 2%       | 2%       | 3%       |
| Prior Arrest                  | 39%      | 42%      | 37%      |

Note: All reported data (except for information on prior arrest) are taken from the baseline interview with the batterers.

to community service, about 77 percent received a conditional discharge, 22 percent received an adjournment in contemplation of dismissal, and 1 percent received traditional probation. For those assigned to the Alternatives to Violence (ATV) treatment groups, about 77 percent received a conditional discharge, 21 percent received an adjournments in contemplation of dismissal, and 2 percent received traditional probation.



Finally, slightly more than one-third (43 percent) of the men said they were married to the victim, another 25 percent reported that the victim was a past intimate partner, and 19 percent reported that they were currently living with, but not married to, the victim. The remaining men reported that they were married, but separated or divorced from the victim (8 percent), or that the victim was a current intimate partner with whom they did not live (9 percent). The remaining 2 percent did not specify a relationship status with the victim. Overall, a typical subject in this sample was a male around age 30, with no prior criminal history, little more than a high school diploma, some but not consistent employment, and a household and personal income of about \$16,000 per year. Typically, he was married to or living with his victim at the time of his arrest.

All the victims were females, with a mean age of 29 years. Six in 10 were African-American (59 percent), 30 percent were Hispanic, 9 percent were white, and 2 percent identified themselves as belonging to some other racial group. The proportion of victims who graduated from high school (66 percent) was comparable to the proportion of high school graduates among batterers. Fewer victims, however, were employed (38 percent), and a large proportion (43 percent) received public assistance. Just 9 percent of the victims reported in the initial interviews that the batterer was their primary source of financial assistance. Victims were poorer than batterers: close to half (46 percent) reported annual household incomes under \$10,000. About two-thirds of the victims and the defendants lived together at the time of arrest (62 percent according to victim interviews, but 70 percent according to batterer interviews). A majority of the victims reported having children with the batterer (79 percent according to victims; 63 percent according to batterers).

The median length of the relationship with the batterer was 5 1/2 years. Victims reported in the initial interview that the violence began, on average, in the second year of the relationship. Sixty-two percent of victims said they had previously called the police because of their perpetrator's abuse, 48 percent had filed a police complaint against their perpetrator in the past, 34 percent reported having had an order of protection against their perpetrator, and 23 percent stated that the perpetrators had been arrested in the past for abusing them. Thus the majority had been abused previously by the batterer. Until the current incident, however, only about one-third had witnessed their batterer formally sanctioned with a restraining order or arrest.

*Assigned Treatments*

The effects of batterer treatment programs can be assessed in several ways. One way is to compare treated groups with those not treated. For example, some judges leave it to probation officers' discretion to assign treatment or not at the time of probation intake. This method is used in Feder's (1996) experiment. That option was not available to us because probation for misdemeanor spouse abuse charges is rare in New York City (only 2 percent of our sample received a regular probation sentence), and judges mandate batterers to treatment; program completion usually is the only condition of plea arrangements. It was not possible to suggest to criminal justice officials that they allow recruited defendants to receive no formal sanctions beyond arrest and conviction. We needed an alternative sanction for the control group, preferably one which does not appear to be related to preventing domestic violence and one without a therapeutic regimen. Community service, as defined below, was such a sanction, and criminal justice officials agreed to use it as an alternative to batterer treatment for men whom we designated as controls. All participants in our experiment were assigned to receive either 40 hours of group batterer treatment or 40 hours of community service.

We evaluated Victim Services' Alternatives to Violence (ATV) program, which operates in New York City and is based on the Duluth model. This model, rooted in a feminist perspective, assumes that domestic violence is a by-product of male and female sex roles, which result in an imbalance of power. The tested curriculum included defining domestic violence, understanding the historical and cultural aspects of domestic abuse, and reviewing criminal/legal issues. Through a combination of instruction and discussion, participants were encouraged to take responsibility for their anger, actions, and reactions. One male and one female leader conducted sessions in either English or Spanish. The program mandated 40 hours of attendance at weekly group meetings over at least eight weeks.

Judges required defendants rejected by lottery from batterer treatment to participate in 40 hours of community service over a two week period.<sup>5</sup> Participants were assigned to renovate housing

---

<sup>5</sup> We found variation in the amount of time participants needed to complete either community service or batterer treatment. Community service usually was completed in about a month; batterer treatment took at least two months. Participants, however, still were under some court supervision before their six-month or one-year sentence expired, and could face a harsher penalty if they violated the terms of their sentence (e.g., a protection order). One could argue that even though the numbers of mandated hours were equal, attending the batterer program over a longer period might make a difference. If this theory were correct, we would expect the difference in recidivism rates between the two groups to dissipate over time.

units, clear vacant lots for community gardens, paint senior citizen centers, and clean up playgrounds; none of these tasks would be expected to affect abusive behavior.

Participants in both batterer treatment and community service programs were expelled from the programs if they developed a pattern of nonattendance: three misses constituted grounds for dismissal from the ATV program. Such cases were referred to the prosecutor's office for action. At the district attorney's discretion, delinquent cases were returned to the court calendar, and new sentences could be imposed.

### *Assignment Process and Case Intake*

Cases were drawn from three of eight postarrestment parts in Kings County Criminal Court. Two of the parts were devoted specifically to domestic violence cases. The third was the jury trial part, to which domestic violence and other cases were transferred if the parties could not reach a negotiated disposition. When the judge, the prosecutor, and the defense agreed the batterer treatment was appropriate, the prosecutor called the ATV office in the court building. The ATV intake person or the research assistant picked up the defendant in court and brought him to the ATV office for an intake interview.

After the defendant completed the interview, his name and case identifier were written on a new line of a logbook. Each line carried a preassigned designation (batterer treatment or community service), set by using a random number table. The use of the log with predetermined treatment assignments and the presence of a research assistant on the three busiest days of the week helped ensure the integrity of the random assignment process. Defendants assigned to batterer treatment were given a start date (usually within a week of intake) and received directions to the class. The defendant was accompanied back to the courtroom, and the prosecutor was told of the lottery assignment. The prosecutor then informed the judge, who in turn disposed the case in keeping with the assignment.

---

To test this possibility we examined a slightly different version of our time-to-failure/survival model (see Table 4). In a model the same as that in Table 4 but with the addition of a time-dependent covariate, this additional variable of time was not significant. Also, in our other prevalence models for officially recorded failures (see Table 4), significant findings are present at six and 12 months postsentencing. If the length of time to complete the program exerted an effect, differences that existed at the six-month follow-up might disappear at 12 months.

*Experimental Misassignments*

Treatment misassignments are a major problem in conducting randomized experiments in the field (Berk, Smyth, and Sherman 1988): that is, cases which the subjects do not receive the treatment or control condition to which they are randomly assigned. The judges in our study could "override" cases when they felt that following the random assignment process could cause egregious harm. Judges overrode 14 percent of our random community service assignments, and instead mandated the ATV program. In no case did the judge override a treatment assignment and reassign the defendant to a control group. Another problem is the situation of those who start a treatment program but do not finish and do not receive full treatment. This occurred in our sample. When either or both these problems arise, one is tempted to make comparisons based on delivered treatment instead of assigned treatment. This solution may appear to be fairer because the treatment of interest will be tested with only those who receive the full dosage.

However, as in procedures used in analyzing the Spouse Assault Replication Program data, we chose to "analyze as randomized" rather than according to the treatment received by perpetrators (see Garner, Fagan, and Maxwell 1995). Analysis of cases according to actual treatment received defeats the purpose of randomizing cases: that is, creating equivalent groups of cases before treatment. In our case, the crossovers were created because judges intervened in the random assignment process. Gartin (1995) makes another argument for analyzing as randomized. In policy studies such as ours, he argues, the issue is not the effect of the treatment per se, but the effect of a policy to apply treatment.

Nevertheless, we considered this a serious issue and performed several tests to assess the extent of possible bias in our results. As one potential test of bias, Sherman (1992) proposed following the "analyze as randomized" dictum as long as the proportion of treatment crossovers does not exceed the proportion of cases with negative outcomes. In our study the crossover rate was 14 ( $n = 53$ ), which did not exceed our one-year combined crime report rate of 17 percent nor the one-year, victim-reported recidivism rate of 19 percent. Berk et al. (1991) recommended conducting tests for any pattern to misassignments, so as to understand what types of subjects the judges may have shifted from the assigned treatments. Similarly, Berk et al. (1988) suggested that researchers generate a statistical model of the misassignment process.

In Table 2, under the heading "Treatment Overridden," we report our effort to predict misassignment through a logistic regression model with five predictor variables (age, ethnicity, marital

status, employment, and prior arrest history). None of the predictor variables approached statistical significance; the percentage of override cases, however, was fairly small (14 percent). Thus, although the judge's abrogation of the random assignment in a minority of cases was not likely to be a random process, we believe that the number of missassignments did not measurably skew the equivalence we had sought through randomization.

### *Interview Procedure*

We attempted to interview victims on three occasions: at case intake or date of court disposition, at six months after intake, and at 12 months after intake. We made the first attempt at all victim interviews by telephone. When telephone attempts failed, interview teams went to victims' homes. If the home interview attempts failed, we mailed letters offering first \$25 and then \$50 for an interview. At the third stage, we also turned over 70 difficult cases to a licensed private investigator; the investigator found five more victims.

We took a number of steps to ensure participants' safety. First, to guarantee confidentiality, the research staff described the project to nonparticipants as a health study. Second, batterers were not present during interviews. The research staff conducted interviews at times that were convenient and safe for the victim, and at local community centers or at Victim Services offices if it was deemed unsafe to conduct the interviews in the victim's home or over the phone. Finally, professional counselors were available to assist the participants if needed.

Our completion rate for the victim surveys was 51 percent for the first interview (not shown in Table 2), 48 percent for the second, and 50 percent for the third. In addition, completion rates by assigned treatment groups did not differ significantly (51 percent vs. 50 percent at time 1; 41 percent vs. 50 percent at time 2; 52 percent vs. 48 percent at time 3). We were unable to contact 131 victims (35 percent) at any time during the follow-up period. In many cases, the victims had moved. Research in other cities (see Davis, Smith and Nickles 1997) with court-involved domestic violence victims has shown that this population is highly transient: many of those who were staying with the batterer or with family members at the time of arrest moved shortly thereafter. Among the victims we located, 7 percent refused to take part or terminated the interview early. Inaccurate or outdated information obtained from prosecutors' files was the primary reason for not conducting interviews with victims.

**Table 2. Implementation of Experimental Design**

|                                          | Treatment Overridden |                 | Completed Victim Interviews |                 |           |                 |
|------------------------------------------|----------------------|-----------------|-----------------------------|-----------------|-----------|-----------------|
|                                          | 14%                  | (53)            | 6 Months                    |                 | 12 Months |                 |
| Base Rate                                |                      |                 | 48%                         | (171)           | 50%       | (186)           |
| Logistic Regression                      | <i>b</i>             | Exp( <i>B</i> ) | <i>b</i>                    | Exp( <i>B</i> ) | <i>b</i>  | Exp( <i>B</i> ) |
| Age                                      | .01                  | 1.01            | -.00                        | 1.00            | -.01      | .99             |
| Ethnicity (African-American)             |                      |                 |                             |                 |           |                 |
| Hispanic                                 | -.28                 | .76             | .54                         | 1.72*           | -.46      | .63             |
| West Indian/Caribbean                    | -.06                 | .95             | -.27                        | .76             | -.46      | .63             |
| Other race                               | -.35                 | .70             | -1.26                       | .29***          | -.78      | .46*            |
| Married                                  | .52                  | 1.68            | .41                         | 1.51            | .22       | 1.25            |
| Employed                                 | -.45                 | .64             | .22                         | 1.24            | -.07      | .93             |
| Number of Prior Arrests                  | .07                  | 1.07            | .02                         | 1.02            | -.01      | .99             |
| ATV Treatment Assigned                   |                      |                 | -.32                        | .73             | .23       | 1.26            |
| Police Recorded Failure Before Interview |                      |                 | -.22                        | .80             | .07       | 1.07            |
| Intercept                                | -2.18***             |                 | -.52                        |                 | .11       |                 |
| Initial Log-Likelihood                   | 305.84               |                 | 518.17                      |                 | 521.24    |                 |
| Final Log-Likelihood                     | 298.96               |                 | 496.23                      |                 | 511.59    |                 |
| P Value                                  | .44                  |                 | .01                         |                 | .38       |                 |

\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ 

We also examined variation in victims' interview completion rates (see Table 2) by the batterer's age, ethnicity, marital status, employment status, and prior arrests. We used batterer characteristics because they were available for virtually the entire sample and because they have been the primary control variables used in other research on interventions to prevent domestic violence. In addition, we examined variation in victim interview rates by assignment to treatment versus control group and by whether the police had recorded any failure before the interview. We found no significant differences in interview completion rates as a function of any these variables, except for ethnicity (see Table 2). We were more successful in interviewing Hispanic victims ( $b = .54$ ;  $p < .05$ ) than African-American victims for the six-month victims interviews, but the completion rate was higher for African-American victims than victims from one of the "other" racial groups (mostly whites and Asians) at both six months ( $b = -1.26$ ;  $p < .001$ ) and 12 months ( $b = -.78$ ;  $p < .05$ ). We also examined the effects, on interview completion rates, of the batterer's income and education, and the relationship of the victim to the batterer. These variables, not shown in Table 2, were not significant.

Reviewers of the batterer treatment literature (e.g., Edleson 1996; Gondolf 1997) have stressed the importance of obtaining high response rates with respondents. Our follow-up attrition rate of

about half the participants clearly falls short of this goal. To investigate the impact of the attrition, we conducted a number of analyses and found no measurable differences between participants and nonparticipants other than the victim's ethnicity. Neither this analysis nor any of our statistical corrections, however, can rule out the possibility of unmeasured differences between these groups, which might interact with treatment assignment to produce a biased comparison test.

### *Measures of Recidivism*

We collected data from several sources to develop multiple indicators of new violence by the batterer against the victim, including arrest reports, crime complaints (which may or may not have resulted in an arrest), and victims' reports of violence by the batterer. These sources have been used in other domestic violence studies that track batterers, such as The Spouse Assault Replication experiments conducted by the National Institute of Justice (see Garner et al. 1995). Other research has shown that it is important to capture outcomes from a variety of sources and along several dimensions because different outcome measures do not always behave similarly (Davis and Taylor 1997). From our data we constructed four recidivism measures: prevalence, rate or frequency of failures, severity, and time to the first failure. Our study included both six- and 12-month postsentencing measures. This combination of data sources, measurement dimensions, and time intervals potentially could produce 24 basic outcomes or recidivism models. Because of limitations in data collection, however, we obtained only enough information to construct 13 outcome models (see Tables 4 to 6).

*Victim self-report surveys.* To assess frequency and severity of violence through victim interviews, we employed Harrell's (1991) revision of the Conflict Tactics Scale (Straus 1979).<sup>6</sup> Harrell's scale measures the frequency of 11 violent acts.<sup>7</sup> In the outcome models, we examined the combined prevalence and frequency of the 11 violent acts and the prevalence and frequency of the six most violent of

<sup>6</sup> This study was limited by budgetary constraints on survey length; thus we could not measure the psychological and emotional abuse that some batterers may have adopted to avoid future arrest.

<sup>7</sup> (1) Forced you to have sex; (2) choked or strangled you; (3) threatened to kill you; (4) beat you up; (5) threatened you with a knife, gun, or other weapon; (6) used a knife, gun, or other weapon against you; (7) threw something at you; (8) pushed, grabbed, or shoved you; (9) slapped or spanked you with an open hand; (10) kicked, bit, or hit you with a fist; and (11) hit or tried to hit you with something.

these acts (our severe victimization measure).<sup>8</sup> The previous two months were the reference period for the scale (as opposed to the previous six months for the criminal justice measures). We reasoned that, if treatment did make a difference, the effect would take some time to appear. Thus, by asking victims to report at the six-month interval about the entire period, we could include reports of violent incidents committed shortly after batterers were assigned to treatment. By deciding on the two-month reference period, we also ensured that any reported violence would have occurred after batterers had completed most of their 40 hours of treatment. Unfortunately, as in other studies, we interviewed victims involved in the triggering court case, but not any new partners the batterers may have started to see.<sup>9</sup>

*Criminal justice records.* We searched computerized records of the Criminal Justice Agency (CJA) and the New York City Police Department (NYPD) to determine whether a new crime report was filed or the batterer was arrested during the study period. We accessed CJA's database of New York City arrests via the court docket numbers of cases. Docket numbers led us to defendants' state criminal identification numbers, which we used to determine whether the defendants were arrested during the 12 months after sentencing. When new incidents were found, we recorded the arrest date and the charge. In addition, we searched the district attorney's computer database, using the docket number to determine whether the victim in the new incident was the same as the victim in the original incident. Because the searches used ID numbers, we are confident that our information on new arrests is highly accurate.

We also searched the NYPD's computerized records to determine whether new crime complaints had been filed against the defendant since sentencing. These searches, conducted by NYPD personnel, used the batterer's name and incident addresses. Therefore the searchers were subject to some error whenever the batterer's name or the street names were spelled incorrectly in the database. Also, each police precinct maintains its own database. When batterers commit a crime outside their home precinct, their home precinct is supposed to receive a record, but we do not know

---

<sup>8</sup> (1) Forced you to have sex; (2) choked or strangled you; (3) threatened to kill you; (4) beat you up; (5) threatened you with a knife, gun, or other weapon; (6) used a knife, gun, or other weapon against you.

<sup>9</sup> The failure to interview all partners may result in an underreporting of domestic violence incidents. Such underreporting, however, should affect only the treatment comparisons if batterers in one of the treatment groups were more likely to find new partners than batterers in the other group. The major obstacle to interviewing these new partners is the great expense of finding them and persuading them to agree to the interview.



how reliably information is transferred. When we found new incidents, officers recorded the date, the nature of the complaint, and whether the complainant was the victim in the original case. As a result of these limitations, we suspect that the NYPD data undercounted victims' complaints. We have no reason to believe, however, that the undercounting would vary according to treatment assignment.

To further reduce measurement error, we merged the CJA arrest reports and the police complaint data into one measure that captured the number of documented criminal justice incidents (e.g., arrests or crime complaints) involving both the defendant and the victim after the treatment was assigned. In this step we followed the procedure used by Maxwell (1998) in the most recent reanalysis of the Spouse Assault Replication Program experiments.

### *Correcting for Missing Information*

Much (if not all) research in behavioral, economic, and social science is plagued by missing information (Winship and Mare 1992). There are two types: item nonresponse and unit nonresponse (Little and Schenker 1995). In the former, missing information is unobserved or unmeasured information on one or more variables for a subset of cases in a database; often it occurs because researchers do not record certain responses or because participants fail to provide certain responses (Weisberg 1985). In the second type of missing information, cases included in a study represent nonrandom samples of a population. Unit nonresponse is often known as sample selection bias (Dubin and Rivers 1989). Unlike the first type, unit nonresponse typically is created when subjects act so as to make it impossible for the researchers to obtain their responses. The nonrandom selection of cases from the entire population into a study is itself a social process and an aspect of social science that is often overlooked (Winship and Mare 1992). Our experiment presented two occasions for sample selection bias: one at the six-month victim interview and another at the 12-month interview.

We addressed both types of missing information. The first problem was missing information for some cases on the nonexperimental covariates. These covariates in the recidivism models were taken from our baseline batterer data set. Although we had conducted baseline interviews with 95 percent of the batterers, not all these men answered all the questions: 22 cases were missing age information; 24 lacked employment status; and 25 were missing ethnicity and marital status. To address this problem, we used imputed quasi-valid values for the missing data, employing a multiple

regression procedure. We constructed a regression model that computed a predicted value for all cases on the basis of cases with valid data, and then used these predicted values to replace the remaining missing data.

To handle sample selection bias or missing victim interviews, we tested a two-step process proposed by Heckman (1979). The first step was to specify a model using a multiple regression of the selection process that would be captured in a single latent measure. We used two different models, one for the six-month interview and another for the 12-month interview (see Table 2). As stated earlier, the batterer's ethnicity was the only statistically significant measure in either the six- or the 12-month victim interviews. We estimated the experiential outcome models with a traditional covariate for sample selection bias included along with the assigned treatment covariate. The addition of this covariate did not change the relationship between the treatment and the outcome measures in any of the experimental models.

To address differences across the treatment groups and potential sample selection bias, we used another method: replacing the sample selection covariate with a set of substantively interesting control measures. Introducing covariates in analyzing data from a randomized experiment is not strictly necessary: randomization should ensure that other (known and unknown) measures that are related to the failure measure, such as the suspect's age or prior criminal record, are distributed similarly across the study groups. In our study, this similarity apparently was achieved. (See our discussion of pretreatment comparisons between the treatment and the control group.)

Nevertheless, introducing covariates is increasingly common in analyzing data from randomized experiments (Patel 1996) for several reasons. First, statistical controls for other factors tend to improve the precision of the treatment comparisons and to correct for any major imbalance in the distribution of these measures across treatments that may have occurred by chance (Armitage 1996). Second, because the batterers assigned randomly to the same treatment group are not exactly alike, statistical controls can address the natural variations between batterers within each study group (Gelber and Zelen 1986). Third, an experimental analysis typically tests only for the average effects of treatment across all batterers, whatever their characteristics. Additional nonexperimental hypotheses, however, can specify other expected direct effects on the outcomes, such as age, and how treatment effects may vary across dimensions of other uncontrolled extraneous factors such as marital status, employment level, or prior criminal record.

The tests for the additional direct effects follow each of the regression models that test for only the direct effects of treatment. We used the following nonexperimental measures in our models: defendant's age, ethnicity, relationship status with the victim, employment level, and prior arrests. All of these measures have been associated with general offending patterns (Blumstein et al. 1986) as well as with violence between intimates (Fagan and Browne 1994; Fagan, Garner, and Maxwell 1997; Hotelling and Sugarman 1986).

## RESULTS

We begin by examining the official criminal justice measures of prevalence (new police complaints or arrests involving the same perpetrator and the same victim) at six months and at 12 months postsentencing. Later we examine our victim-based measure of prevalence — the prevalence of new reports of violence made by victims during research interviews — at the six month and the 12-month follow-up postsentencing. We conducted the prevalence comparisons using logistic regression.

In the next set of analyses we examine incidence/frequency of failure, using a count model (negative binomial regression) that tests whether the distribution of failures (cases in which new violence occurred) differed according to treatment. We used negative binomial regression to model the distribution that we observed: a large majority of the sample did not fail at all during the time observed, some failed once, fewer failed twice, and only a handful failed more often. This kind of highly skewed distribution violates the normality assumption of ordinary least squares, even with log or other data transformations.

In the final set of analyses, using proportional hazard models, we examined differences between treatment conditions in time elapsed to first failure. Using a time-to-failure model, we could address the question of timing not only at six and at 12 months, but at any point during the follow-up period. We have only one time-to-failure model based on the official recidivism measures, because we did not collect all of the information during the victim interviews that we needed to construct this measure using the data reported by victims.

In addition, for the final three tables (4, 5 and 6) we estimate two models. The top rows in each table present the experimental model, or our type 1 models. These initial models specify only the treatment assignment measure. The second set of rows presents the full model, or our type 2 models. These final models specify a

more complete set of structure measures, along with the treatment assignment measure.

Table 3 presents the distribution of officially recorded offenses and the victim-reported incidents by assigned treatment. As reported there, the failure rate in the ATV treatment group is significantly smaller (about 50 percent), based on officially recorded incidents, at six months ( $M = .38$  for the controls, and  $M = .16$  for the treatment group;  $f = 8.09$ ;  $p < .01$ ) and again at 12 months ( $M = .55$  for controls, and  $M = .28$  for treatment group;  $f = 6.82$ ;  $p < .01$ ). On the basis of victims' reports, four out of four comparisons again show lower failure rates among the ATV treatment group; not one of the four rate differences was significant, however. Thus, on the basis of these six comparisons, the treatment group showed positive change (e.g., reduction in the mean number of failures), but only two of the six positive effects reach statistical significance. In the remaining analysis we focus more closely on these experimental comparisons by using appropriate regression techniques, modeling different outcome dimensions, and controlling statistically for sample differences.

**Table 3. Frequency of Outcome Incidents, by Assigned Treatment**

|            | Officially Recorded Incidents |     |           |     | Victim-Reported Incidents |     |         |     |           |     |         |     |
|------------|-------------------------------|-----|-----------|-----|---------------------------|-----|---------|-----|-----------|-----|---------|-----|
|            | 6 Months                      |     | 12 Months |     | 6 Months                  |     |         |     | 12 Months |     |         |     |
|            |                               |     |           |     | Any                       |     | Severe  |     | Any       |     | Severe  |     |
|            | %                             | %   | %         | %   | %                         | %   | %       | %   | %         | %   | %       | %   |
|            | Control                       | ATV | Control   | ATV | Control                   | ATV | Control | ATV | Control   | ATV | Control | ATV |
| 0          | 79                            | 90  | 74        | 84  | 79                        | 78  | 80      | 86  | 78        | 85  | 86      | 93  |
| 1          | 12                            | 6   | 13        | 10  | 4                         | 8   | 3       | 6   | 7         | 7   | 6       | 2   |
| 2          | 4                             | 3   | 7         | 5   | 5                         | 6   | 9       | 3   | 4         | 4   | 1       | 2   |
| 3          | 4                             | 1   | 3         | 1   |                           |     |         |     |           |     |         |     |
| 4          | 1                             | 1   | 2         | 1   |                           |     |         |     |           |     |         |     |
| 5          | 1                             |     | 1         |     | 7                         | 5   | 4       | 3   | 8         | 3   | 4       | 2   |
| 6          |                               |     | 1         |     |                           |     |         |     |           |     |         |     |
| 7          |                               |     | 1         | 1   |                           |     |         |     |           |     |         |     |
| 8          |                               |     |           |     | 5                         | 3   | 4       | 3   | 1         |     | 1       |     |
| 9          |                               |     |           |     |                           |     |         |     |           |     |         |     |
| 10+        |                               |     |           |     |                           |     |         |     | 2         | 1   | 2       | 1   |
| Mean       | .38                           | .16 | .55       | .28 | .90                       | .67 | .76     | .45 | .99       | .46 | .74     | .32 |
| N          | 186                           | 190 | 186       | 190 | 93                        | 78  | 93      | 78  | 90        | 99  | 90      | 99  |
| Chi-Square | 11.10*                        |     | 8.97      |     | 1.86                      |     | 4.50    |     | 3.86      |     | 4.53    |     |
| F-Test     | 8.09**                        |     | 6.82**    |     | 2.37                      |     | 4.20    |     | 12.96     |     | 8.36    |     |

\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

### *Criminal Justice Recorded Failures*

*Prevalence of official failures.* The first results presented in Table 4 detail the prevalence of officially recorded failures based on a logistical regression using the classic experimental analysis (our type 1 models). That is, we compare the proportion of the treatment group

failing with the proportion of the control group failing, without controlling for other independent variables. Our six-month type 1 model shows, as expected, that treatment assignment was associated with a lower prevalence of officially recorded failures ( $b = -.97$ ); this difference is statistically significant with a two-tailed test ( $p < .001$ ). The difference or effect size is also substantively large: the model estimated that 59 percent fewer than the expected number of subjects assigned to the ATV treatment group committed a recorded offense against the victim ( $\text{Exp}(B) = .38$ ).

The full model (type 2) for the six-month results builds on this first model by adding additional control measures to account for the natural heterogeneity between and within the two comparison groups. Again, in terms of the proportion of officially recorded failures, we found a statistically significant difference between the treatment group and the control group ( $b = -.86$ ;  $p < .01$ ). The sizable reduction in the proportion of batterers with any incident within six months remains the same in the experimental group. No other batterer characteristics were related significantly to a change in the prevalence of officially recorded complaints.

In addition, whether a victim interview was completed before we searched the criminal history files was not statistically significant in this six-month model (nor in any of the models reported in Tables 4, 5, and 6). Some observers voiced concerns that interviewed victims may constitute a different kind of sample, or that interviews could lead to more violence; thus we were pleased to find no significant effect from the interview measure. Therefore this set of nonsignificant findings (in Tables 4 to 6) tends to support the idea that the victims who completed an interview were not significantly different, in terms of victimization risk, from those who did not.

In the 12-month follow-up using the experimental model, we found results nearly identical to those of the six-month model. This later model showed again that treatment was associated with a significantly lower prevalence of officially recorded failures ( $b = -.61$ ,  $p < .01$ ). At 12 months, however, the difference in incident rates between the two experimental groups had diminished (from  $\text{Exp}(B) = .41$  to  $\text{Exp}(B) = .55$ ). This suggests that the early effect of treatment may lessen with time. Similar results from the type 2 model for the 12-month failure measure showed a statistically significant difference between the treatment group and the control group ( $b = -.59$ ;  $p < .05$ ). In another parallel with the six-month analyses, none of the additional control measures introduced in the type 2 model were significantly correlated with the likelihood of any failure within 12 months.

**Table 4. Officially Recorded Incidents Since Assigned Treatment**

|                              | Prevalence <sup>a</sup> |         |           |         | Rate <sup>b</sup> |         |           |         | Time to Failure <sup>c</sup> |        |
|------------------------------|-------------------------|---------|-----------|---------|-------------------|---------|-----------|---------|------------------------------|--------|
|                              | 6 Months                |         | 12 Months |         | 6 Months          |         | 12 Months |         | 12 Months                    |        |
|                              | b                       | Exp(B)  | b         | Exp(B)  | b                 | Exp(B)  | b         | Exp(B)  | b                            | Exp(B) |
| Type 1                       |                         |         |           |         |                   |         |           |         |                              |        |
| ATV (assigned)               | -.97                    | .38***  | -.61      | .55*    | -.84              | .43**   | -.70      | .50*    | -.55                         | .58*   |
| Alpha                        |                         |         |           |         | -.98***           |         | -.43      |         |                              |        |
| Intercept                    | -1.29***                |         | -1.03***  |         | 4.24***           |         | 3.46***   |         |                              |        |
| Initial Log-Likelihood       |                         | -163.38 |           | -194.62 |                   | -262.99 |           | -350.09 |                              | 924.42 |
| Final Log-Likelihood         |                         | -158.59 |           | -191.78 |                   | -229.07 |           | -295.26 |                              | 918.54 |
| P value                      |                         | .00     |           | .02     |                   | .00     |           | .00     |                              | .02    |
| Type 2                       |                         |         |           |         |                   |         |           |         |                              |        |
| ATV (assigned)               | -.86                    | .42**   | -.59      | .55*    | -.84              | .43*    | -.70      | .50*    | -.53                         | .59    |
| Prior victim interview       | -.02                    | .98     | -.19      | .83     | .13               | 1.14    | .02       | 1.02    | -.20                         | .81    |
| Age                          | .02                     | 1.02    | .01       | 1.01    | .00               | 1.00    | .00       | 1.00    | .01                          | 1.01   |
| Ethnicity (African-American) |                         |         |           |         |                   |         |           |         |                              |        |
| Hispanic                     | -.61                    | .54     | -.33      | .72     | .56               | 1.75    | -.41      | .66     | -.35                         | .70    |
| West Indian/Caribbean        | -.68                    | .51     | -.64      | .53     | -.66              | .52     | -.64      | .53     | -.55                         | .58    |
| Other race                   | -.75                    | .47     | .78       | 2.18    | -.27              | .76     | -.33      | .72     | -.70                         | .50    |
| Married                      | .23                     | 1.26    | .12       | 1.13    | .34               | 1.40    | .28       | 1.32    | .12                          | 1.13   |
| Employed                     | -.54                    | .58     | -.28      | .76     | -.25              | .78     | -.27      | .76     | -.28                         | .76    |
| Prior arrest                 | .25                     | 1.28    | .36       | 1.43    | .22               | 1.25    | .18       | 1.20    | .38                          | 1.46   |
| Alpha                        |                         |         |           |         | -.93              |         | -.43      |         |                              |        |
| Intercept                    | -1.46                   |         | -.94      |         | 3.74**            |         | 3.46***   |         |                              |        |
| Initial Log-Likelihood       |                         | -163.38 |           | -194.56 |                   | -254.68 |           | -350.09 |                              | 924.42 |
| Final Log-Likelihood         |                         | -152.03 |           | -186.56 |                   | -225.51 |           | -295.26 |                              | 906.10 |
| P value                      |                         | .01     |           | .06     |                   | .00     |           | .00     |                              | .03    |

<sup>a</sup>Logistic Regression<sup>b</sup>Negative Binomial Regression<sup>c</sup>Cox Regression\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ 

*Frequency of official failure.* As in the earlier logistical regression, the type 1 or experimental model of the frequency of failure at the six-month follow-up showed results that were both in the expected direction (treatment lowers frequency of failures) and statistically significant ( $b = -.84$ ;  $p < .01$ ). Once again, the difference was also substantively large. Over the first six months, the batterers assigned to the treatment group experienced 57 percent fewer incidents than the control group ( $\text{Exp}(B) = .43$ ). The six-month type 2 model (see Table 4) similarly showed a statistically significant difference between the treatment group and the control group in the frequency of officially recorded failures ( $b = -.84$ ;  $p < .05$ ). Also, as in the early logistic regression models, no other significant control variables correlated with a change in the frequency of officially recorded complaints.

Corresponding results also emerged for the 12-month count models (see Table 4). The type 1 model for the 12-month results showed that treatment is associated with a significantly lower frequency of officially recorded failures ( $b = -.70$ ;  $p = .05$ ). As in the prevalence models, however, the size of the effect was diminished somewhat by 12 months: batterers assigned to the treatment group experienced a 50 percent reduction in incidence compared with the control group ( $\text{Exp}(B) = .50$ ), as opposed to a 57 percent reduction at

six months. Finally, the type 2 model for the 12-month results continued to show a statistically significant difference between the treatment group and the control group in the frequency of officially recorded failures ( $b = -.70$ ;  $p < .05$ ). Once again, none of the covariates were statistically significant.

*Time to first official failure.* Our final analyses using the official data, reported in Table 4, examine the timing of the first officially recorded incident after the treatment assignment. These results are based on a Cox regression that modeled the hazard of time to first new officially recorded offense if it occurred within the first 12 months of follow-up. As in the earlier results, the type 1 model showed a significant decrease in the hazard rate for the treatment group compared with the rate for the control group ( $b = -.55$ ;  $p < .05$ ). In other words, the batterers assigned to the treatment group registered a 42 percent reduction in the likelihood of a new incident on any given day after the treatment assignment ( $\text{Exp}(B) = .58$ ).<sup>10</sup> The type 2 model also showed a statistically significant difference between the treatment group and the control group in time to first failure ( $b = -.53$ ;  $p < .05$ ). Once again, the control measures were not related significantly to the hazard rate.

Overall, the analysis based only on the official recorded data reveals consistently less recidivism among batterers assigned to the ATV treatment group. Batterer treatment seemed not only to reduce the prevalence and frequency of officially recorded recidivism over 12 months, but also to create a consistent period of greater safety for victims during the first year of follow-up. In other words, the positive treatment effect found at six months did not diminish significantly at 12 months. Also, the batterer treatment was not trivial in reducing the number of incidents between the batterer and the victim known to the police. Depending on the measure and the length of follow-up, the effect of the ATV treatment varies between 41 percent and 59 percent less recidivism than would be expected. Across the 10 models based on criminal justice records, the weighted average was 50 percent less than expected.

#### *Incidents Reported by Victims to Research Interviewers*

Early in Table 3 we saw that the treatment and the control groups had virtually the same prevalence and incident rates for

<sup>10</sup> We conducted a Cox regression with a treatment by time-dependent interaction term to test for the proportionality of the hazard rates over the first 12 months of follow-up. The coefficient for the interaction term was about .0001 and was non-significant in both the experimental and the full models. Therefore we removed the interaction term from the analysis to simplify the interpretation and presentation of the results.

"any" victimization as reported by the victims at both six months (treatment = 78 percent no victimization, .67 mean; controls = 79 percent no victimization, .90 mean) and 12-months (treatment = 85 percent no victimization, .46 mean; controls = 78 percent no victimization, 0.99 mean). We also found similar nonsignificant differences between the control and the experimental groups for the "severe" victimization measure (see Table 3). In this final set of analyses we examine the victim interview data, using a similar set of regression models to those used for the official recorded failures.

*Six-month treatment effects.* Table 5 presents results of logistic and negative binomial regression for the prevalence and frequency of victim-reported failures (any victimization or severe victimization) by the six-month follow-up interview. The results from the type 1 model, which estimated the prevalence of any victimization, showed that the effect of the treatment measure was slightly positive but not significant. The type 2 model, with the additional control measures, also showed that the treatment variable was nonsignificant, although the estimated reduction in the prevalence of victimization was about 16 percent. Yet in contrast to the earlier models based on the official data, with the type 2 model we found that prior arrest was statistically associated with a rather sizable increase in the risk of any repeat victimization ( $b = 1.03$ ;  $p < .01$ ). The batterers with a prior arrest history were 2.8 times more likely to fail on this measure than batterers without any such history.

With regard to severe victimization, the type 1 and type 2 models in Table 5 produce results similar but not identical to those found in the "any victimization" models. The treatment assignment measure was still nonsignificant in both models, but in both instances the effect was negative and the reduction was not small. In the full model, the reduction in prevalence was estimated to be slightly more than 50 percent of the base rate. In addition, the results from the full model for severe victimization showed not only that prior arrest was related significantly to an increased risk of victimization ( $b = 1.15$ ;  $p < .01$ ), but also that Hispanics were 3.4 times more likely than African-Americans ( $b = 1.21$ ;  $p < .05$ ) to severely assault their intimate partner.

Table 5 also presents the victim-reported frequency (negative binomial) of "any victimization" and "severe victimization" at six months. The two type 1 models show that treatment assignment was associated with a lower frequency of victim reports of both "any victimization" ( $b = -.30$ ) and "severe victimization" ( $b = -.53$ ); neither of the two coefficients was statistically significant, however. We found nearly identical results for the two full models. Findings



**Table 5. Victims' Reports of Incidents at Six Months  
(n = 171)**

|                              | Prevalence <sup>a</sup> |        |        |        | Frequency <sup>b</sup> |         |        |         |
|------------------------------|-------------------------|--------|--------|--------|------------------------|---------|--------|---------|
|                              | Any                     |        | Severe |        | Any                    |         | Severe |         |
|                              | b                       | Exp(B) | b      | Exp(B) | b                      | Exp(B)  | b      | Exp(B)  |
| Type 1                       |                         |        |        |        |                        |         |        |         |
| ATV (assigned)               | .02                     | 1.02   | -.45   | .64    | -.30                   | .74     | -.53   | .59     |
| Alpha                        |                         |        |        |        | 7.94                   | .00***  | 9.85   |         |
| Intercept                    | -1.29***                |        | -1.36  | .00*** | -.10                   | .00     | -.27   |         |
| Initial Log-Likelihood       |                         | -89.31 |        | -79.81 |                        | -294.67 |        | -349.92 |
| Final Log-Likelihood         |                         | -89.31 |        | -78.81 |                        | -170.97 |        | -156.05 |
| P value                      |                         | .96    |        | .28    |                        | .00     |        | .00     |
| Type 2                       |                         |        |        |        |                        |         |        |         |
| ATV (assigned)               | -.17                    | .84    | -.75   | .47    | -.69                   | .50     | -.52   | .59     |
| Prior victim interview       | -.54                    | .59    | -.34   | .71    | -.51                   | .60     | -.19   | .83     |
| Age                          | .04                     | 1.04   | .04    | 1.04   | .06                    | 1.06    | .07    | 1.08    |
| Ethnicity (African-American) |                         |        |        |        |                        |         |        |         |
| Hispanic                     | .75                     | 2.11   | 1.21   | 3.36*  | 1.01                   | 2.75    | 1.42   | 4.13    |
| West Indian/Caribbean        | -.03                    | .97    | .24    | 1.27   | -.58                   | .58     | -.42   | .65     |
| Other race                   | 1.09                    | 2.97   | 1.36   | 3.88   | .63                    | 1.87    | .47    | 1.60    |
| Married                      | -.66                    | .52    | -.70   | .49    | -1.38                  | .25     | -1.61  | .20     |
| Employed                     | -.25                    | .78    | -.32   | .73    | .19                    | 1.21    | -.00   | 1.00    |
| Prior arrest                 | 1.03                    | 2.80** | 1.15   | 3.17** | .90                    | 2.45    | .61    | 1.85    |
| Alpha                        |                         |        |        |        | 6.30                   | .00***  | 7.97   | .00***  |
| Intercept                    | -2.46                   | .00*** | -3.08  | .00**  | -2.05                  | .00     | -2.64  | .00     |
| Initial Log-Likelihood       |                         | -89.31 |        | -79.41 |                        | -269.99 |        | -234.44 |
| Final Log-Likelihood         |                         | -89.31 |        | -69.43 |                        | -164.75 |        | -138.72 |
| P value                      |                         | .96    |        | .02    |                        | .00     |        | .00     |

<sup>a</sup>Logistic regression<sup>b</sup>Negative binomial regression

\*p &lt; .05; \*\*p &lt; .01; \*\*\*p &lt; .001

for the treatment assignment measures for both "any" and "severe" victimization were nonsignificant. Nevertheless, in both models, the reduction in the expected frequency of victimization estimated by the full model was not small. For "any victimization," the expected frequency of failure among those assigned to the treatment was reduced by about 50 percent. The reduction was about 40 percent for "severe victimization."

*Twelve-month treatment effects.* Table 6 displays the 12-month results for the prevalence and incidence of victim-reported incidents ("any" and "severe" victimization). Essentially these 12-month results mirror the six-month findings: we found all negative relationships between treatment assignment and victimization rates, but all still were nonsignificant. Two control measures in the full models were significant, however. The prevalence models showed that for batterers married to their victim, the risk of any victimization one year later was reduced significantly. Similarly, the risk of severe victimization was lower when the batterer was employed at the time of the treatment assignment. These two significant effects, however, did not remain in the frequency models.

As with the official criminal justice measures, consistent results emerged. For 15 of the 16 victim-based recidivism models, the

**Table 6. Victims' Reports of Incidents at 12 Months  
(n = 189)**

|                              | Prevalence <sup>a</sup> |        |          |        | Frequency <sup>b</sup> |         |         |         |
|------------------------------|-------------------------|--------|----------|--------|------------------------|---------|---------|---------|
|                              | Any                     |        | Severe   |        | Any                    |         | Severe  |         |
|                              | b                       | Exp(B) | b        | Exp(B) | b                      | Exp(B)  | b       | Exp(B)  |
| Type 1                       |                         |        |          |        |                        |         |         |         |
| ATV (assigned)               | -.47                    | 0.63   | -.80     | .45    | -.76                   | .47     | -.83    | .43     |
| Alpha                        |                         |        |          |        | 9.87***                |         | 21.75** |         |
| Intercept                    | -1.25***                |        | -1.78*** |        | -.11                   |         | .30     |         |
| Initial Log-Likelihood       |                         | -90.56 |          | -63.82 |                        | -327.30 |         | -288.34 |
| Final log-likelihood         |                         | -89.78 |          | -62.46 |                        | -164.79 |         | -111.98 |
| P value                      |                         | .21    |          | .10    |                        | .00     |         | .00     |
| Type 2                       |                         |        |          |        |                        |         |         |         |
| ATV (assigned)               | -.43                    | .65    | -.71     | .49    | -.83                   | .44     | -.61    | .54     |
| Prior victim interview       | .95                     | 2.58   | -.05     | .95    | .32                    | 1.38    | -.54    | .58     |
| Age                          | .01                     | 1.01   | -.01     | .99    | .02                    | 1.02    | .01     | 1.01    |
| Ethnicity (African-American) |                         |        |          |        |                        |         |         |         |
| Hispanic                     | .06                     | 1.07   | -.15     | .86    | -.81                   | .44     | -.67    | .51     |
| West Indian/Caribbean        | -.02                    | .98    | .48      | 1.62   | .58                    | 1.78    | 2.48    | 11.98   |
| Other race                   | .42                     | 1.52   | -.60     | .55    | -.54                   | .58     | -1.02   | .36     |
| Married                      | -1.07                   | .34**  | -.85     | .43    | -.60                   | .55     | -.71    | .49     |
| Employed                     | -.34                    | .71    | -1.09    | .34*   | -.60                   | .55     | -2.05   | .13     |
| Prior arrest                 | -.70                    | .50    | -.48     | .62    | -.57                   | .56     | -.63    | .53     |
| Alpha                        |                         |        |          |        | 8.56***                |         | 13.67** |         |
| Intercept                    | -1.48                   |        | -.52     | .00    | .06                    |         | .49     |         |
| Initial Log-Likelihood       |                         | -90.56 |          | -63.82 |                        | -307.20 |         | -242.60 |
| Final Log-Likelihood         |                         | -84.11 |          | -57.78 |                        | -161.18 |         | -104.98 |
| P value                      |                         | .17    |          | .15    |                        | .00     |         | .00     |

<sup>a</sup>Logistic regression<sup>b</sup>Negative binomial regression\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ 

estimated effect for treatment was negative (although the results were nonsignificant); collectively these 15 negative effects suggest that persons assigned to treatment may report 44 percent fewer than the expected incidents of victimization. Even though none of the 16 tests achieved statistical significance, the empirical regularity of the results suggests that treatment is working.

## DISCUSSION

A number of reviewers (Hamberger and Hastings 1993; Rosenfeld 1992; Tolman and Bennett 1990) have observed that after years of asking "Are batterer treatment programs effective?" we still do not have a conclusive answer. Yet a new generation of evaluations (based on experimental models) promises to provide such an answer. Results from our experiment are consistent with those obtained by Palmer et al. (1992). Both experiments found that the rate of new incidents reported to criminal justice authorities was reduced significantly among batterers assigned to treatment.

As in the Palmer et al. (1992) study, our treatment effects were relatively large. In the analyses of the prevalence and frequency rates, we found that men assigned to a therapeutic treatment group

for batterers were less likely to be subject to future criminal complaints and arrests involving the same victim than were men assigned to a control group: the rate was 59 percent smaller at the six-month follow-up and 45 percent smaller at 12 months. Also, the 12-month follow-up showed that men assigned to batterer treatment, on average, registered a longer period of nonviolence or longer time to first failure than men assigned to the control group. On the basis of the victim survey data, the size of the treatment effect was similar to that found in the criminal justice data: the average violence rate was 44 percent smaller as a result of treatment. The effects were nearly always in the same direction. In 15 of 16 models, victims of men assigned to the treatment group reported fewer violent acts than victims of men in the control group.

We are more cautious about the outcomes of victim surveys, however, because in contrast to the results based on the criminal justice records) none of the differences between the two groups were statistically significant at the standard .05 alpha level. This non-significance may be due in part to the small sample sizes ( $n = 171$  for the six-month data;  $n = 189$  for the 12-month data). Alternatively, as one anonymous peer reviewer suggested, the men assigned to batterer treatment may have become more skillful than their control group counterparts at avoiding calls to the police when they are violent.

Some observers are likely to question the generalizability of our positive treatment findings because our sample represented only a small proportion of the spouse abuse cases adjudicated by the Brooklyn Criminal Court during our intake period: 3.4 percent, or 376 of nearly 11,000 adjudicated cases, met our study criteria. The sample size was limited because the prosecutor, the defendant, and the judge had to agree to treatment in order to make a case eligible for the lottery process. Therefore the treatment effects we observed cannot be generalized even to the population of adjudicated spouse abuse cases in the Brooklyn Criminal Court. Similarly, many of the earlier quasi-experimental evaluations of batterer treatment programs can be criticized on the grounds that participants are only a small and probably unrepresentative proportion of batterers processed in the courts from which their samples were drawn. Thus the Palmer et al. (1992) experiment is probably the most readily generalizable because it used a more representative sample of all batterers placed on probation by the court where the study was conducted. Their sample size, however, was only 59 cases, and not all of those batterers were necessarily motivated to undergo treatment.

Which of the two sampling approaches is better for evaluation? The answer depends on public policies defining who should participate in treatment programs. Many program administrators argue that treatment is appropriate only for batterers who demonstrate a willingness to change their behavior. We accept that argument; accordingly our sampling approach makes more sense because it filtered potentially unmotivated batterers. The Palmer et al. (1992) study, however, provides a test of cases more typically referred to court-mandated batterer groups.

The issue of generalizability crops up in another way as well. In many treatment studies that have relied on batterer or victim surveys to assess violence, interview response rates have been poor, some as low as 30 percent. Low response rates are a problem because the cases in which follow-up data are available may be different from those in which data are not available. For example, Edleson and Syers (1990) reported higher levels of education and income for batterers who completed follow-up surveys than for those who did not. Therefore it is unclear whether their analysis of treatment effects applies to the low-SES batterers who did not complete the survey as well as the higher-SES batterers who did so.

Although we succeeded in interviewing only about half of the victims in our sample, we were able to demonstrate the statistical equivalence of completers and noncompleters across a range of participant characteristics. Moreover, we had complete official recidivism data on virtually the entire sample, and we found that rates of victimization reported to the police did not differ significantly between victims who completed the surveys and those who did not.

Another limitation of our study is the problem of treatment misassignment: cases assigned to the control group that a judge reassigned to the treatment group. Our study includes 53 misassigned cases; this raises some questions about the precision of our treatment estimates. Consequently our results may reflect more clearly the potential for failures under a policy of assigning people to treatment than the theoretical issue of how well treatment works for those who receive it. Probably this misassignment tends to suppress the effectiveness of the ATV treatment program which we estimated. That is, if the 53 batterers in our control group who benefited from treatment because of the misassignment had been denied treatment, the effect of treatment would have been increased. Therefore it is likely that our reported treatment estimates are conservative.

Studies of batterer treatment programs have grown increasingly sophisticated: They use designs with a high degree of internal validity. As recognized by Fagan (1996), randomized experiments

are the preferred evaluation design. The few experimental evaluations that have been completed provide useful information on statistical power and sample sizes for designers of future studies. As more studies are completed, we may learn with some confidence how much and under what conditions treatment programs reduce violence.

If that proves to be the case, attention will turn to policy issues regarding how these programs are conducted. For example, the trend has been to lengthen treatment; yet in contrast to findings in the therapeutic drug treatment literature (see S.R. Maxwell 1994), virtually no research supports one treatment length over another. Also, practitioners in the field disagree about the content of treatment sessions. Many programs have adopted a feminist orientation which assumes that eliminating violent behavior requires changing the participant's perception of men's and women's roles in society. Other programs take a different approach, emphasizing anger control, stress management, and better communication techniques. Currently there is no empirical basis for deciding which approach is best. One thing seems certain, however: group treatment programs for batterers will continue in some form, if only because no good sentencing alternatives exist for spouse abusers.

## REFERENCES

- Armitage, P. 1996. "The Design and Analysis for Clinical Trials." Pp. 1-30 in *Design and Analysis of Experiments: Handbook of Statistics*, vol. 13, edited by S. Ghosh and C. R. Rao. Amsterdam: Elsevier.
- Berk, R.A., H. Black, J. Lilly, and G. Rikoski. 1991. *Evaluating Alternative Police Response to Spouse Assault in Colorado Springs: An Enhanced Replication of the Minneapolis Experiment, 1987-1989*. Colorado Springs, CO: Colorado Springs Police Department.
- Berk, R.A., G.K. Smyth, and L.W. Sherman. 1988. "When Random Assignment Fails: Some Lessons from the Minneapolis Spouse Abuse Experiment." *Journal of Quantitative Criminology* 4:209-33.
- Blumstein, A., J. Cohen, J. Roth, and C. Visher, eds. 1986. *Criminal Careers and Career Criminals*. Washington, D.C.: National Academy Press.
- Buzawa, E.S. and C.G. Buzawa. 1996. *Domestic Violence: The Criminal Justice Response*. 2nd ed. Newbury Park, CA: Sage.
- Chen, H.T., C. Bersani, S. Myers, and R. Denton. 1989. "Evaluating the Effectiveness of a Court Sponsored Abuser Treatment Program." *Journal of Family Violence* 4:309-22.
- Crowell, N.A. and A.W. Burgess, eds. 1996. *Understanding Violence Against Women*. Washington, D.C.: National Academy Press.
- Davis, R.C., B.E. Smith, and L. Nickles. 1997. *Prosecuting Domestic Violence Cases With Reluctant Victims: Assessing Two Novel Approaches*. Washington, DC: American Bar Association.
- Davis, R.C. and B.G. Taylor. 1997. "A Proactive Response to Family Violence: The Results of a Randomized Experiment." *Criminology* 35:307-33.
- . 1999. "Does Batterer Treatment Reduce Violence? A Synthesis of the Literature." *Women and Criminal Justice* 10(2):69-93.

- Dobash, R., R. Dobash, K. Cavenagh, and R. Lewis. 1995. "Evaluating Criminal Justice Programmes for Violent Men." Pp. 88-116 in *Gender and Crime*, edited by R. Dobash and R. Dobash. Cardiff: University of Wales Press.
- . 1996. *Research Evaluation of Programmes for Violent Men*. Edinburgh: Centre Research, Scottish Office.
- Dubin, J.A. and D. Rivers. 1989. "Selection Bias in Linear Regression, Logit and Probit Model." *Sociological Methods and Research* 18(2):360-90.
- Dunford, F. 1999. "The San Diego Navy Experiment: An Assessment of Interventions for Men Who Assault Their Wives." Unpublished Manuscript.
- Dutton, D. 1986. "The Outcome of Court-Mandated Treatment for Wife Assault: A Quasi-Experimental Evaluation." *Violence and Victims* 1(3):163-75.
- . 1988. "Profiling of Wife Assaulters: Preliminary Evidence for a Trimodal Analysis." *Violence and Victims* 3(3):5-29.
- Edleson, J. 1996. "Controversy and Change in Batterers' Programs." Pp. 154-69 in *Future Interventions with Battered Women and Their Families*, edited by J. Edleson and Z. Eisikovits. Thousand Oaks, CA: Sage.
- Edleson, J. and M. Syers. 1990. "Relative Effectiveness of Group Treatments for Men Who Batter." *Social Work Research and Abstracts* 26(2):10-17.
- Eisikovits, Z. and J. Edleson. 1989. "Intervening with Men Who Batter: A Critical Review of the Literature." *Social Service Review* 37:84-414.
- Fagan, J.A. 1996. "The Criminalization of Domestic Violence: Promises and Limits." *NIJ Research Report*. Washington, DC: National Institute of Justice, U.S. Department of Justice.
- Fagan, J.A. and A. Browne. 1994. "Violence Against Spouses and Intimates." Pp. 112-51 in *Understanding and Controlling Violence*, vol. 3., edited by A.J. Reiss Jr. and J.A. Roth. Washington, DC: National Academy Press.
- Fagan, J., E. Friedman, S. Wexler, and V.L. Lewis. 1984. *National Family Violence Evaluation*. San Francisco, CA: URSA Institute.
- Fagan, J.A., J. Garner, and C.D. Maxwell. 1997. *Reducing Injuries to Women in Domestic Assaults*. Atlanta, GA: Centers for Disease Control and Prevention.
- Feder, L. 1996. *A Test of the Efficacy of Court-Mandated Counseling for Domestic Violence: A Broward County Experiment*. Proposal submitted to the National Institute of Justice. Florida Atlantic University, Boca Raton, Florida.
- . 1999. "The Efficacy of Court-Mandated Counseling for Convicted Misdemeanor Domestic Violence Offenders: The Broward Experiment." Washington, DC: American Society of Criminology.
- Garner, J., J.A. Fagan, and C.D. Maxwell. 1995. "Published Findings From the Spouse Assault Replication Program: A Critical Review." *Journal of Quantitative Criminology* 11:3-28.
- Gartin, P.R. 1995. "Dealing with Design Failures in Randomized Field Experiments: Analytic Issues Regarding the Evaluation of Treatment Effects." *Journal of Research in Crime and Delinquency* 32:425-45.
- Gelber, R.D. and M. Zelen. 1986. "Planning and Reporting of Clinical Trials." Pp. 406-25 in *Medical Oncology*, edited by P. Calabresi, P.S. Schein, and S.A. Rosenberg. New York: Macmillan.
- Gondolf, E. 1995. *Batterer Intervention: What We Know and Need to Know*. Washington, DC: National Institute of Justice.
- . 1997. "Expanding Batterer Program Evaluations." Pp. 208-18 in *Out of Darkness: Contemporary Research Perspectives on Family Violence*, edited by G. K. Kaufman and J. Jasinski. Thousand Oaks, CA: Sage.
- Greenfeld, L.A., M.R. Rand, D. Craven, P. Klaus, C.A. Perkins, C. Ringel, G. Warchol, C. Maston, and J. Fox. 1998. *Violence by Intimates: Analysis of Data on Crimes by Current or Former Spouses, Boyfriends, and Girlfriends*. Washington, DC: U.S. Department of Justice.
- Hamberger, K. and J. Hastings. 1993. "Court-Mandated Treatment of Men Who Assault Their Partner: Issues, Controversies and Outcomes." Pp. 188-232 in *Legal Responses to Wife Assault: Current Trends and Evaluation*, edited by Z. Hilton. Newbury Park, CA: Sage.
- Hanna, C. 1996. "No Right to Choose: Mandated Victim Participation in Domestic Violence Prosecutions." *Harvard Law Review* 109:1849-1910.
- Harrell, A. 1991. *Evaluation of Court-Ordered Treatment of Domestic Violence Offenders*. Washington, DC: Urban Institute.

- Healey, K., C. Smith, and C. O'Sullivan. 1998. *Batterer Intervention: Program Approaches and Criminal Justice Strategies*. Washington, DC: U.S. Department of Justice.
- Heckman, J.J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1):153-61.
- Hotaling, G.T. and D. Sugarman. 1986. "An Analysis of Risk Markers in Husband to Wife Violence: The Current State of Knowledge." *Violence and Victims* 1(2):101-24.
- Little, R.J.A. and N. Schenker. 1995. "Missing Data." Pp. 39-76 in *Handbook of Statistical Modeling for the Social and Behavioral Sciences*, edited by G. Arminger, C.C. Clogg, and M.E. Sobel. New York: Plenum.
- Maxwell, C.D. 1998. "The Specific Deterrent Effect of Arrest on Aggression Between Intimates and Spouses." PhD dissertation, Graduate School of Newark, School of Criminal Justice, Rutgers University.
- Maxwell, S.R. 1994. "Formal and Informal Social Controls of Drug Treatment Retention of Offenders." PhD dissertation, Graduate School of Newark, School of Criminal Justice, Rutgers University.
- Palmer, S., R. Brown, and M. Barrera. 1992. "Group Treatment Program for Abusive Husbands: Long-Term Evaluation." *American Journal of Orthopsychiatry* 62: 276-83.
- Patel, H.I. 1996. "Clinical Trials in Drug Development: Some Statistical Issues." Pp. 31-57 in *Design and Analysis of Experiments: Handbook of Statistics*, vol. 13, edited by S. Ghosh and C.R. Rao. Amsterdam: Elsevier.
- Rebovich, D.J. 1996. "Prosecution Response to Domestic Violence: Results for a Survey of Large Jurisdictions." Pp. 175-76 in *Do Arrest and Restraining Orders Work?*, edited by E.S. Buzawa and C.G. Buzawa. Thousand Oaks, CA: Sage.
- Rosenbaum, A. and K.D. O'Leary. 1986. "The Treatment of Marital Violence." Pp. 44-69 in *Clinical Handbook of Marital Therapy*, edited by N.S. Jacoben and A.S. Gurmand. New York: Guilford.
- Rosenfeld, B.D. 1992. "Court Ordered Treatment of Spouse Abuse." *Clinical Psychological Review* 12:205-26.
- Saunders, D. 1996. "Feminist-Cognitive-Behavioral and Process-Psychodynamic Treatments for Men Who Batter: Interaction of Abuser Traits and Treatment Models." *Violence and Victims* 11:393-414.
- Saunders, D.G. and S.T. Azar. 1989. "Treatment Programs for Family Violence." Pp. 481-546 in *Family Violence*, vol. 11, edited by L. Ohlin and M. Tonry. Chicago, IL: University of Chicago Press.
- Sherman, L.W. 1992. *Policing Domestic Violence: Experiments and Dilemmas*. New York: Free Press.
- Straus, M.A. 1979. "Measuring Family Conflict and Violence: The Conflict Tactics Scale." *Journal of Marriage and the Family* 41:75-88.
- Tjaden, P. and N. Thoennes. 1998. *Prevalence, Incidence and Consequences of Violence Against Women: Findings from the National Violence Against Women Survey*. Washington, DC: National Institute of Justice.
- Tolman, R.M. and L. Bennett. 1990. "A Review of Quantitative Research on Men Who Batter." *Journal of Interpersonal Violence* 5(1):87-118.
- Tolman, R.M. and J.L. Edelson. 1995. "Interventions for Men Who Batter: A Review of Research." Pp. 211-55 in *Understanding Partner Violence: Prevalence, Causes, Consequences, and Solutions*, edited by S.M. Stith and M.A. Straus. Minneapolis, MN: National Council on Family Relations.
- Weisberg, S. 1985. *Applied Linear Regression*. New York: Wiley.
- Winship, C. and R.D. Mare. 1992. "Models for Sample Selection Bias." *Annual Review of Sociology* 11(1):111-44.