The Effectiveness of Batterer Intervention Programs

A Literature Review
&
Recommendations for Next Steps

[Abridged version without copies of articles]

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

March 2011

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

TABLE OF CONTENTS

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 <u>not</u> 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, and Appendix D for summaries of mid-quality studies.

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 <u>more</u> 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.

Part 4: Conclusions and Recommendations for Next Steps

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

CDC

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, PhD[1]

Kathleen C. Basile, PhD[2]

Marci F. Hertz, MS[2]

Dylan Sitterle, BS[1]

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

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

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 ☐

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.

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.

_____ 2. I push and shove my partner around violently.

_____ 3. I hit and punch my partner's arms and body.

_____ 4. I threaten my partner with a weapon.

_____ 5. I beat my partner so hard he or she must seek medical help.

_____ 6. I slap my partner around his or her face and head.

_____ 7. I beat my partner when I'm drinking.

_____ 8. I make my partner afraid for his or her life.

_____ 9. I physically throw my partner around the room.

____ 10. I hit and punch my partner's face and head.

____ 11. I beat my partner in the face so that he or she is ashamed to be seen in public.

____ 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.

____ 15. I knock my partner down and then kick or stomp him or her.

____ 16. I twist my partner's fingers, arms or legs.

____ 17. I throw dangerous objects at my partner.

____ 18. I bite or scratch my partner so badly that he or she bleeds or has bruises.

____ 19. I violently pinch or twist my partner's skin.

____ 20. I hurt my partner while we are having sex.

____ 21. I hurt my partner's breast or genitals.

____ 22. I try to suffocate my partner with pillows, towels, or other objects.

____ 23. I poke or jab my partner with pointed objects.

____ 24. I have broken one or more of my partner's bones.

____ 25. I kick my partner's face and head.

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

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

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

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 ☐

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.

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

	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.

_____ 2. I expect my partner to obey.

_____ 3. I become very upset and angry if my partner says that I have been drinking too much.

_____ 4. I demand my partner to perform sex acts that he or she does not enjoy or like.

_____ 5. I become very upset if my partner's work is not done when I think it should be.

_____ 6. I don't want my partner to have any male friends.

_____ 7. I tell my partner he or she is ugly and unattractive.

_____ 8. I tell my partner to hop to it when I give him or her an order.

_____ 9. I expect my partner to hop to it when I give him or her an order.

____ 10. I insult or shame my partner in front of others.

____ 11. I become angry if my partner disagrees with my point of view.

____ 12. I carefully control the money I give my partner.

____ 13. I tell my partner that he or she is dumb or stupid.

____ 14. I demand that my partner stay home.

____ 15. I don't want my partner to work or go to school.

____ 16. I don't want my partner socializing with his or her female friends.

____ 17. I demand sex whether my partner wants it or not.

____ 18. I scream and yell at my partner.

____ 19. I shout and scream at my partner when I'm drinking.

____ 20. I order my partner around.

____ 21. I have no respect for my partner's feelings.

____ 22. I act like a bully towards my partner.

____ 23. I frighten my partner.

____ 24. I treat my partner like he or she is a dimwit.

____ 25. I'm rude to my partner.

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	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 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

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 batter 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, El, 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

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 <u>all</u> 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, the 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. *Violence and Victims,* 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. *International Journal of Offender Therapy and Comparative Criminology,* 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 on 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. & 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.

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.