

Methodological Review of Intimate Partner Violence Prevention Research

By: [Christine E. Murray](#) and Jennifer Graybeal

Murray, C. E., & Graybeal, J. (2007). Methodological review of intimate partner violence prevention research. *Journal of Interpersonal Violence*, 22(10: October 2007), 1250-1269.

Made available courtesy of Sage Publications: <http://www.sagepub.com/>

*****Reprinted with permission. No further reproduction is authorized without written permission from Sage Publications. This version of the document is not the version of record. Figures and/or pictures may be missing from this format of the document.*****

Abstract:

The authors present a methodological review of empirical program evaluation research in the area of intimate partner violence prevention. The authors adapted and utilized criterion-based rating forms to standardize the evaluation of the methodological strengths and weaknesses of each study. The findings indicate that the limited amount of empirical research related to intimate partner violence prevention demonstrates some consistent methodological strengths and limitations. The authors conclude with a presentation of recommendations for further research in the area of intimate partner violence prevention.

Keywords: domestic violence; intimate partner violence; methodological review; violence research

Article:

Intimate partner violence (IPV) is a major public health concern that produces significant costs for victims, perpetrators, family members, law enforcement agencies, health care organizations, and the general population (Coker, 2004; Lee, Sanders Thompson, & Mechanic, 2002; Osofsky, 2003; Plichta, 2004). Because of the costs associated with IPV, researchers and advocates have begun to focus attention on prevention (Coker, 2004; Fullwood, 2002; Mitchell-Clark & Autry, 2004). Several preventive initiatives have been implemented in communities across the United States (Fullwood, 2002; Mitchell-Clark & Autry, 2004). However, research in this area is limited (Guterman, 2004; Hickman, Jaycox, & Aronoff, 2004; Wolfe & Jaffe, 1999).

This article discusses the methodological strengths and limitations of research in the area of IPV prevention. We conducted a methodological review of empirical research based on the procedures followed by Heneghan, Horwitz, and Leventhal (1996) in their review of family preservation programs. For the current review, we included empirical research in the category of program evaluation research. Program evaluation studies are used to determine the outcomes of a particular program as delivered to a target population (Leber, St. Peters, & Markman, 1996).

We included program evaluation studies covering the three categories of preventive interventions: primary, secondary, and tertiary (Coker, 2004). Primary prevention efforts target the general population and are designed to prevent IPV from ever occurring. Examples of primary prevention programs include media campaigns and public policy efforts to change social norms that promote violence. Secondary prevention efforts involve early detection and intervention of IPV to prevent violence from developing into a more severe problem. An example of a secondary prevention initiative is routine screening for IPV by health care professionals. Finally, tertiary prevention initiatives attempt to prevent death or disability due to current violence. These efforts aim to prevent future violence, with the assumption that violence is an existing problem in the target population. Examples of tertiary prevention initiatives include battered women's shelters and some batterer intervention programs.

Review of the Literature

In this section, we describe the contribution that systematic methodological reviews can make to a body of literature. We then review literature describing the state of research in this area.

Overview of Systematic Methodological Reviews

The American Educational Research Association (2006) defined *methodological reviews* as “descriptions of research design, methods, and procedures that ... highlight the strengths and weaknesses of methodological tools and explore how methods constrain or open up opportunities for learning about educational problems” (¶ 6). Lilford et al. (2001) suggested that systematic methodological reviews should be based on broad-based searches of literature from a wide range of sources. These reviews involve several stages: planning, gathering the studies to include, analyzing the methodologies used in the studies, and synthesizing the information gained through the review. Systematic methodological reviews have the potential to inform practice and provide evidence to support ongoing practices (Gallagher, 1999). In sum, systematic reviews are used to provide a synthesis of the existing literature in an area of research and to generate new ideas and directions for practice (Dickersin, 2002).

In the area of research on IPV, methodological reviews have been conducted to examine whether the existing research supports the practice of universal screening for IPV among patients in health care settings. In 2004, the U.S. Preventive Services Task Force (USPSTF) made public the results of their systematic review of a large body of research on this topic. Their conclusion was that there was not sufficient evidence to make a recommendation either for or against universal screening for family violence. In addition, Ramsay, Richardson, Carter, Davidson, and Feder (2002) conducted a systematic review of research available in three academic databases. They concluded that the existing research did not provide ample justification to support universal screening for IPV in health care settings. Together, these studies synthesized the existing research related to IPV screening and provided justification for additional research in this area. The current review aims to provide similar information related to IPV prevention program evaluation research.

The State of IPV Prevention Research

Guterman (2004) wrote that a gap exists between research and practice in the area of interpersonal violence prevention. Indeed, many violence prevention programs are used with minimal evidence supporting their effectiveness, and methodological flaws are common in this area of research (Guterman, 2004; Hickman et al., 2004). To establish the need for a systematic review of the methodologies used in this area of research, in this section we review scholarly commentaries on the state of interpersonal violence prevention research.

Researchers who study IPV prevention face numerous practical and ethical challenges. These challenges include stricter requirements for Institutional Review Boards and confidentiality issues surrounding mandatory reporting of child maltreatment (Hickman et al., 2004). In addition, Mitchell-Clark and Autry (2004) wrote that other barriers include the complexity of community-based work, lack of funding for evaluation, and limited experience in evaluation by practitioners.

In addition to practical and ethical challenges, significant methodological challenges are common in IPV prevention research. These limitations include (a) varying definitions of key constructs, (b) measurement issues, (c) a lack of long-term follow-up, and (d) validity issues.

Definitions of key terms vary widely among researchers who study various aspects of family violence (Wallace, 2005). There are no universally accepted definitions of the terms *intimate partner*, *intimate partner violence*, *violence*, and *domestic violence*. These definitional issues translate into measurement issues in that it becomes challenging to measure a construct when no clear, operational definition of that construct exists. For this reason, it is common to observe some studies that define IPV as solely physical violence and other studies in which the definitions of IPV are inclusive of physical, sexual, and emotional violence.

Measurement issues present a number of methodological challenges, particularly in an overreliance on self-report measures (Guterman, 2004)—which may be hindered by social desirability biases—and an inability to measure accurately certain outcome variables using empirical instrumentation (Mitchell-Clark & Autry, 2004). For example, subtle changes in attitudes toward violence, relationships, and sex-role expectations may not be measured easily using self-report instruments. DeVoe and Kantor (2002) also confirmed the importance of selecting psychometrically sound measurement instruments in the area of IPV prevention research.

Guterman (2004) indicated that another area of methodological challenges surrounds the lack of longitudinal research and adequate follow-up assessments. The addition of long-term follow-up assessments produces additional expenses and time requirements for researchers, and resources may not be available to support such efforts. In addition, researchers may experience difficulties in maintaining current contact information for participants during the follow-up period. Despite the challenges presented by long-term research studies, follow-up studies are essential for determining whether observed outcomes of program participation are able to be maintained.

Finally, IPV prevention program evaluation researchers face numerous threats to the internal and external validity of their studies. Potential threats to internal validity include natural maturation, possibly nonequivalent groups prior to the intervention, and attrition over time (Leber et al., 1996). These threats can be addressed by including control groups, using statistical analyses to examine the similarity of the groups prior to the intervention, and keeping track of the reasons that participants withdraw from the study. Challenges related to external validity involve designing studies for which the results will be able to be generalized to the larger population and other settings (Leber et al., 1996). A particularly relevant issue for IPV prevention researchers relates to developing and evaluating interventions that are appropriate for the needs of diverse populations. Lee et al. (2002) described the pressing need for increased research on experiences related to IPV among women of color—particularly in the area of prevention and intervention program evaluation research. The Centers for Disease Control and Prevention (CDC; 2000) funded 10 demonstration projects focusing on IPV prevention among minority populations in an effort to address this need.

The need for additional, methodologically sound research in the area of IPV prevention is clear (Mitchell-Clark & Autry, 2004; Wolfe & Jaffe, 1999). The existing research provides a foundation on which future researchers can strive to build and improve. The scholars cited in this section provide important directions for researchers in the area of IPV prevention. However, they do not provide specific details regarding the methodologies used in existing studies in this area. Therefore, this article addresses a gap in the literature by providing a detailed, systematic methodological review of research related to IPV prevention.

Method

We based our methods for evaluating existing IPV prevention research on the model of Heneghan et al.'s (1996) methodological review of family preservation program research. In this section, we describe the procedures used for finding and selecting studies, our evaluation criteria, and the procedures we used to rate the included studies.

Study Selection Criteria

We used five criteria to select program evaluation research studies to review. Studies were only included if they (a) were related explicitly to the topic of IPV prevention (i.e., the term *prevention* was included in either the title, abstract, or keyword of the article, and a primary purpose of the article was to inform research and/or practice in the area of IPV prevention); (b) described and evaluated a preventive intervention; (c) were available in peer-reviewed journals; (d) demonstrated an empirical basis; and (e) were dated 1990 or later. We excluded studies examining acquaintance rape prevention programs to include only studies that addressed violence within intimate relationships. We included studies related to dating violence prevention.

Studies were located through extensive searches of 12 academic databases in fields associated with IPV, including nursing, counseling, public health, social work, education, women's studies, and psychology. The databases searched were Academic Search Elite, Academic Search Premier, AgeLine, CINAHL, Contemporary Women's Issues, ERIC, Health Source: Nursing/Academic Edition, GLBT Life, MEDLINE, NASW Clinical Register, PsychInfo, and Social Work Abstracts. We used a variety of related search terms to capture the maximum number of possible studies. These search terms included *violence*, *intimate partner violence*, *domestic violence*, *abuse*, *spouse abuse*, *IPV*, *prevention*, *prevention program*, *program evaluation*, *community participatory research*, and *community outreach*.

The abstracts of each study identified through this search were screened to exclude studies not meeting the criteria described above. The full text of each article that appeared to meet the criteria was obtained and reviewed further to exclude those studies not meeting the criteria. The remaining studies were retained for this review.

Evaluation Criteria

To standardize our methodological review procedures, we developed standardized evaluation rating criteria (see the appendix). We adapted these criteria from Heneghan et al. (1996), whose criteria were based on Chalmers et al. (1981). Heneghan et al. developed a 15-item questionnaire to evaluate the most significant methodological characteristics of the studies they reviewed. This questionnaire was specific to research on family preservation programs, so the criteria required modification to be appropriate for the studies we reviewed on IPV prevention. In making these adaptations, we drew on Guterman (2004) and the program evaluation research guidelines of Leber et al. (1996).

Fifteen items made up the research evaluation criteria questionnaire. The following questions were included: (a) Are representative sampling procedures (e.g., random selection, systematic sampling) used to select individual participants? (b) Are eligibility criteria for participation in the study specified? (c) Are exclusion criteria specified? (d) Is there a control group? (e) Are participants randomly assigned to groups at the individual level? (f) Do the researchers document the group similarity (e.g., demographic variables, other relevant variables) prior to the intervention? (g) Is the type of intervention specified for the treatment group? (h) Is the intensity and duration of the intervention specified for the treatment group? (i) Is the type of intervention specified for the control group? (j) Is the intensity and duration of the intervention specified for the control group? (k) Are standardized, psychometrically sound pretest assessment instruments utilized? (l) Are standardized, psychometrically sound outcome measures used? (m) Are outcomes measured at multiple levels (attitudinal, behavioral, and observational)? (n) Are appropriate statistical analyses used (sensitive to measure change over time, as in repeated measures ANOVA or regression)? and (o) Is a follow-up assessment included (at least 1 month following the intervention)? Studies received one point for each question that was answered affirmatively, yielding a total possible score for each study of 15.

We adopted the score stratification guidelines used by Heneghan et al. (1996). Studies receiving scores of at least 70% (11 or more of 15) were considered acceptable studies; studies receiving scores between 40% and 69% (6 to 10 of 15) were considered adequate studies, and studies receiving scores below 40% (0 to 5 of 15) were considered unacceptable.

Rating Procedures

Two forms were constructed: (a) the Methodological Evaluation Criteria Rating Guide (see the appendix) and (b) the Score Form on which scores were recorded. The Rating Guide included additional clarification about scoring each item. The Rating Guide also specifies the original source of each criterion. All studies were reviewed independently by both authors. The Rating Guide demonstrated 78% interrater agreement with a κ of .67 ($p = .000$). Like Heneghan et al. (1996), we discussed score disagreements, referred to the articles to locate the disputed information in text, and reached a consensus score for the items in question. We report consensus scores in the Results section.

Results

Based on our eligibility criteria, we located and reviewed nine program evaluation studies. Table 1 presents the reference, target population, and variables studied for each of the studies we reviewed.

Methodological Review of Program Evaluation Studies

Table 2 presents the nine program evaluation study ratings. Based on the 15-item program evaluation research evaluation criteria questionnaire, one study was deemed acceptable with a score of at least 70% (McFarlane, Soeken, & Wiist, 2000), three studies were deemed adequate with scores between 40% and 69% (Foshee et al., 1998; Macgowan, 1997; Weisz & Black, 2001), and five studies were deemed unacceptable with scores below

40% (Avery-Leaf, Cascardi, O'Leary, & Cano, 1997; Jaffe, Suderman, Reitzel, & Killip, 1992; Lavoie, Vezina, Piche, & Boivin, 1995; Matthews, 2000; Rynerson & Fishel, 1993).

Table 1
Reference, Target Population, and
Variables Assessed in Studies Reviewed

Reference	Target Population	Outcomes/Variables Assessed
Primary and/or secondary prevention program evaluation studies		
1. Avery-Leaf, Cascardi, O'Leary, & Cano (1997)	High school students in New York	Physical aggression, victimization, dating violence, violence attitudes, justification of dating violence
2. Foshee, Bauman, Arriaga, Helms, Koch, & Linder (1998) ^a	8th- and 9th-grade students in North Carolina	Violence victimization and perpetration, acceptance of dating violence norms, gender stereotyping, conflict management, and anger
3. Jaffe, Suderman, Reitzel, & Killip (1992)	High school students in London, Canada	Knowledge about IPV, attitudes toward dating violence, behavioral intentions
4. Lavoie, Vezina, Piche, & Boivin (1995)	10th-grade students in Quebec City, Canada	Knowledge and attitudes toward dating violence
5. Macgowan (1997)	Middle school students	Knowledge about IPV, attitudes about different types of IPV, and attitudes about how to deal with IPV
6. Matthews (2000)	Participants in a violence prevention program	Memories about the program, perception of the impact of the program, importance of program topics
7. McFarlane, Soeken, & Wiist (2000) ^b	Pregnant Hispanic women	Experiences of threatened & actual violence, use of community resources
8. Weisz & Black (2001)	African American middle school students	Knowledge about and attitudes toward dating violence, violence victimization and perpetration

Table 1 (continued)

Reference	Target Population	Outcomes/Variables Assessed
Tertiary prevention program evaluation studies		
9. Rynerson & Fishel (1993)	Male abusers and some of their partners	Locus of control, dyadic adjustment

Notes: IPV = intimate partner violence.

a. Additional follow-up information on this study is available in Foshee et al. (2000) and Foshee et al. (2004). Only the Foshee et al. (1998) study was reviewed for this article as it contained the original intervention outcome study.

b. This study could be classified as either secondary or tertiary prevention, depending on the frequency, duration, and severity of the violence that the female participants had been experienced at the time of the screening for participation in the study and intervention.

The first six research evaluation criteria related to sampling and group assignment procedures. One study used representative sampling procedures. Eight of nine studies described the eligibility criteria for participation in the study. One study specified exclusion criteria that were used to screen out inappropriate participants from the

study. Six studies involved a control group. None of the studies utilized random assignment of participants to groups at the individual level. Six studies documented the similarity of the treatment and control groups prior to the intervention.

The next four criteria related to treatment and control group conditions. Seven studies specified the type of intervention received by the treatment group in sufficient detail. Five studies described the intensity and duration of the intervention condition. Five studies specified the type of intervention received by the control group in sufficient detail. Four studies described the intensity and duration of the control group condition. The next three research evaluation criteria related to measurement procedures. One study used standardized, psychometrically sound pretest instrumentation. One study used standardized, psychometrically sound outcome measures. Three studies measured outcomes at multiple levels (e.g., attitudinal, behavioral, and/or observational).

Table 2
Methodological Criteria Ratings

Criteria	Study										% Meeting Criteria	
	Avery-Leaf	Foshee	Jaffe	Lavoie	Macgowan	Matthews	McFarlane	Weisz	Rynerson			
Representative sampling procedures	0	1	0	0	0	0	0	0	0	0	0	11
Eligibility criteria	1	1	1	1	1	1	1	1	1	0	0	89
Exclusion criteria	0	0	0	0	1	0	0	0	0	0	0	11
Control group	1	1	0	1	1	0	1	1	1	0	0	67
Random assignment	0	0	0	0	0	0	0	0	0	0	0	0
Group similarity	1	1	0	1	1	0	1	1	1	0	0	67
Treatment intervention type	0	1	1	0	1	1	1	1	1	1	1	78
Treatment intervention timing	0	0	1	0	1	0	1	1	1	1	1	56
Control intervention type	1	1	0	0	1	0	1	1	1	0	0	56
Control intervention timing	0	1	0	0	1	0	1	1	1	0	0	44
Pretest assessment instrument	0	0	0	0	0	0	1	0	0	0	0	11
Outcome measures	0	0	0	0	0	0	1	0	0	0	0	11
Multiple levels for outcomes	0	1	0	0	0	0	1	1	1	0	0	33
Statistical analyses	1	1	1	1	1	0	1	1	1	0	0	78
Follow-up assessment	0	1	1	0	0	0	1	1	1	0	0	44
Total Score	5	10	5	4	9	2	12	10	2	10	2	

Note: 1 = present, 0 = absent.

The final two criteria related to the statistical analyses and follow-up assessment. Seven studies demonstrated appropriate use of statistical analytic techniques that are sensitive to measure change over time (e.g., repeated measures ANOVA or regression). Four studies included a follow-up assessment that occurred at least 1 month following the intervention. For one study (Foshee et al., 1998), reports of follow-up assessments were available elsewhere (Foshee et al., 2004; Foshee et al., 2000). Thus, although this article only reported one postintervention assessment 1 month following the intervention, this study received credit for this criterion.

Discussion

Overall, our review demonstrates that scant attention has been paid to empirical research in the area of IPV prevention. Our ability to identify only nine studies is evidence of the limited nature of research in this area. Thus, we view this area of research as being only in the early stages of development. We begin this discussion with an overview of the limitations of this review. Next, we present a summary of the strengths and weaknesses of existing research in the area of IPV prevention. We conclude with a presentation of recommendations for future research based on the findings of this review.

Limitations of this Review

The findings of this review must be considered within the context of three limitations. First, we limited our study to empirical research, thereby excluding qualitative research. The inclusion of only empirical research allowed us to use a systematic evaluation strategy. Future research is needed to examine the merits of existing qualitative research related to IPV prevention.

Second, the criterion-based rating procedure that we adapted from the work of Heneghan et al. (1996) demonstrates two potential limitations. First, this procedure assigns equal weight to each of the criteria on which a study is evaluated. Certain criteria may have a greater impact than other criteria on the overall strength of a study. However, like Heneghan et al. (1996), we believe that our criteria “are minimally necessary for scientific validity, namely, proper assignment to groups, a standardized intervention, and appropriate selection and measurement of outcomes” (p. 540). A second limitation of the adaptation of the Heneghan et al. (1996) procedure relates to the score stratification system according to percentage of criteria met by each study. For example, this stratification system considers studies meeting 40% to 67% of the criteria to be an adequate study. However, some researchers may consider a study meeting only 40% of the criteria to be less than adequate. Therefore, we encourage readers to consider the raw scores and percentages along with its corresponding classification in evaluating each study.

Finally, we were unable to track and report the total number of studies that were excluded from this review because of the procedures we used to locate eligible studies. There was a large amount of overlap in the articles that were located through the 12 academic databases we searched. The abstract and citation of each article was screened immediately through the academic databases, and a large number of studies were screened out immediately on evidence that they did not meet inclusion criteria. In addition, a number of studies were located that were not related directly to the topics of IPV and/or prevention. Therefore, a report on the number of studies for which the abstracts and citations were screened and excluded would not represent an accurate figure of the amount of studies relevant to the topic of IPV prevention research. Therefore, we reported the number of included studies only.

Strengths and Weaknesses of Existing Research

A consideration of the strengths and weaknesses of existing research in the area of IPV prevention can help to inform future research. Regarding the program evaluation research studies we reviewed, the primary strengths included specification of eligibility criteria, provision of adequate description of the content and format of the treatment group interventions, the use of appropriate statistical analyses, and the provision of documentation of the similarity of the treatment and control groups. The limitations of the research included limited use of representative sampling procedures to select individual participants, a lack of specifying exclusion criteria, nonuse of random assignment to groups on an individual level, and not using psychometrically sound assessment instrumentation.

Recommendations for Future Research

We begin this section with an acknowledgment that researchers in the area of IPV prevention face many challenges inherent in the process of studying a complex, sensitive social issue. Despite recent commitments to fund preventive evaluation research by federal funding agencies (CDC, 2000), financial resources for the evaluation of preventive interventions are often in short supply. We agree with other scholars that research in this area is sorely needed (Coker, 2004; Hickman et al., 2004; Lee et al., 2002; Wolfe & Jaffe, 1999), and we encourage scholars and practitioners to continue to advocate for increased financial resources from federal and private funding sources to support research in this area. At this time, many important questions related to IPV prevention remain unanswered (Guterman, 2004). We conclude this methodological review with some important directions for future research.

Sampling procedures and group assignment. We urge researchers to pay increased attention to the methods they use to select study participants and to assign participants to conditions. Only one study we reviewed used representative sampling procedures, and none of the studies used random assignment to groups on an individual basis. The use of these procedures presents practical challenges such as identifying a meaningful population, increased costs associated with more complex research designs, and working within organizations in which groups of participants are already established (e.g., school systems). However, these procedures are important for improving the internal and external validity of study findings. As Christensen and Heavey (1999) wrote,

Selection is a greater problem for prevention than for therapy research because of the broader reach of prevention.... Because these complex issues have the potential to systematically bias the results of prevention studies, they require greater attention than they have received. (p. 184)

Researchers can begin to address the practical issues associated with representative sampling and random assignment to groups by placing meaningful parameters on the populations they choose to study (e.g., first-year college students at small liberal arts colleges in one state), conducting small-scale studies with a strong methodological foundation, and conducting prevention programs outside the realm of preexisting groups (e.g., after school or during lunch time in school systems).

Exclusion criteria. A related issue is the failure of all but one of the studies we reviewed to specify whether exclusion criteria were used to prohibit certain individuals from study participation. Because preventive interventions are developed in accordance with the population for which they are targeted (Coker, 2004), researchers should pay increased attention to specifying the criteria that would be used to exclude inappropriate participants from study involvement. For example, it may be inappropriate, and perhaps even dangerous, for an individual who has already demonstrated a consistent history of violent behavior in intimate relationships to participate in a primary prevention program. Therefore, screening guidelines should be used and explicated in research reports. In addition, researchers can assist potential program adopters by providing guidelines for removing program participants if it becomes apparent that they are not appropriate recipients of the intervention once the program has started (e.g., a participant who makes inappropriate, offensive jokes about interpersonal violence throughout the program or a participant who makes threatening comments to another participant).

Assessment instrumentation. Only one of the studies we reviewed demonstrated the use of standardized, psychometrically sound assessment instrumentation. Most studies used instruments that the researchers either developed themselves or modified from existing instruments. Several studies provided psychometric properties of their instrumentation; however, these properties generally did not meet our evaluation criteria. Researchers have an interest in using measures that are targeted to their program and study content. However, an overreliance on inadequate measurement techniques renders research data that is unclear and difficult to compare across studies (Gay & Airasian, 2000). For research to advance in the area of IPV prevention, researchers must select instrumentation very carefully and ensure that the measures they use demonstrate adequate psychometric properties. When program evaluators must utilize locally developed instrumentation, the psychometric properties of these instruments should be presented in the study findings. In addition, evaluations using locally developed instrumentation will be strengthened through inclusion of existing instruments that

measure similar constructs as a validity check. We acknowledge that the existing instrumentation may be irrelevant or inadequate for specific populations and interventions. For this reason, we also encourage increased dialogue between researchers and practitioners to ensure that researchers are aware of the measurement needs of practitioners in the field.

Multiple levels of assessment. Only three of the nine studies we reviewed employed assessment strategies that examined multiple levels of variables. Most studies relied solely on attitudinal measures. Although attitudes and beliefs certainly influence individuals' experiences with violent behaviors, these attitudes may not translate into behavioral change. As Nabi, Southwell, and Hornik (2002) wrote, "General and specific beliefs about domestic violence, though perhaps predictive of intentions to act in response to a domestic violence situation, may not predict actual behaviors in the face of a real domestic abuse incident" (p. 443). Therefore, researchers should examine violence perpetration and victimization experiences in addition to attitudinal variables. We also urge researchers to use additional objective forms of assessment, such as observational coding of behaviors, rather than rely solely on self-report measures (Guterman, 2004).

Scope of research studies. Finally, we encourage researchers in the area of IPV prevention to expand the scope of the interventions they study. Most studies we reviewed focused on interventions targeted to individuals, and all studies primarily assessed outcomes at the individual level. This sole focus on individuals ignores the powerful influence of the larger social context on experiences of IPV. Therefore, researchers should develop strategies that combine individually based programs with programs that address larger social networks. One study we examined (Foshee et al., 1998) provides a model of an intervention that encompassed the individual, organizational, and community levels. Other researchers (Budde & Schene, 2004; Sabol, Coulton, & Korbin, 2004) confirmed the importance of engaging communities and social networks to effectively prevent violence. As Gil (1996) discussed, prevention efforts are futile if they do not actively aim to change the elements of the larger social structure that contribute to violence (e.g., economic conditions, social norms, and oppression of groups of people).

Appendix: Rating Guide

1. *Are representative sampling procedures used to select individual participants (e.g., random selection, systematic sampling)?* (Leber, St. Peters, & Markman, 1996) One point will be awarded if you answer *yes* to either of the following questions: (a) Are participants randomly selected into the sample? and (b) Is a systematic sampling procedure used (e.g., systematic sampling)?
2. *Are eligibility criteria for participation in the study specified?* (Heneghan, Horwitz, & Leventhal, 1996) One point will be awarded if there are clearly stated eligibility criteria specified in the text of the article.
3. *Are exclusion criteria specified?* (Heneghan et al., 1996) One point will be awarded if the researchers describe criteria for which potential participants would not be eligible to participate in this study. **NOTE:** Exclusion criteria are different than eligibility criteria (though these are related). Eligibility criteria specify the general parameters that define the population/sample (e.g., all college males age 18 – 24 years who are members of a fraternity). Exclusion criteria specify what members of that population would not be appropriate to be in the study (e.g., had already demonstrated an act of domestic violence, had never been in a romantic relationship). It is possible for an article to specify one of these types of criteria but not the other. If there are no exclusion criteria, that should be noted.
4. *Is there a control group?* (Heneghan et al., 1996) One point will be awarded if a control group of any type is involved.
5. *Are participants randomly assigned to groups?* (Heneghan et al., 1996) One point will be awarded if participants are randomly assigned to either the treatment or control group on an individual basis. For example, a study in which participants are "randomly" assigned to a treatment or control condition based on which class

they are in at school would not receive a point for this criterion. (No point shall be awarded for this category if no control group exists.)

6. *Do the researchers document the “similarity of the groups with regard to sociodemographic characteristics” prior to the intervention?* (Heneghan et al., 1996, p. 539) One point will be awarded if the researchers report a comparison of the treatment and control groups based on relevant demographic characteristics (e.g., gender, age, education, having witnessed violence in one’s family-of-origin). (No point shall be awarded for this category if no control group exists.)

7. *Is the type of intervention specified for the treatment group?* (Heneghan et al., 1996) One point will be awarded when the intervention is described sufficiently that it would allow replication, particularly related to an adequate description of the content and format of the intervention condition.

8. *Is the intensity and duration of the intervention specified for the treatment group?* (Heneghan et al., 1996) One point will be awarded if factors related to the timing and duration of the intervention are described adequately. A point should only be awarded if the researchers specify the number of meetings, the length of each meeting, and the frequency in which meetings occurred.

9. *Is the type of intervention specified for the control group?* (Heneghan et al., 1996) One point will be awarded when the control condition is described sufficiently that it would allow replication, particularly related to an adequate description of the content and format of the control condition. (No point shall be awarded for this category if no control group exists.)

10. *Is the intensity and duration of the intervention specified for the control group?* (Heneghan et al., 1996) One point will be awarded if factors related to the timing and duration of the control group are described adequately. A point should only be awarded if the researchers specify the number of meetings, the length of each meeting, and the frequency in which meetings occurred. (No point shall be awarded for this category if no control group exists.)

11. *Are standardized, psychometrically sound (acceptable reliability and validity described in the text) pretest assessment instruments utilized?* (Leber et al., 1996) One point will be awarded when the pretest instrumentation involves psychometrically sound instruments that have been validated and shown to be reliable. Basic psychometric properties must be presented in the text of this article to receive this point. (No point shall be awarded if there is no pretest assessment.) NOTE: Psychometrically sound instruments will demonstrate the following properties (Corcoran & Fischer, 2000): (a) internal consistency (Cronbach’s alpha) of at least 0.80; (b) test–retest reliability of at least $r = .80$ (especially when measuring change over time as in pretest–posttest evaluation); and (c) construct validity (some evidence should be presented to demonstrate that the instrument correlates with other measures of similar constructs).

12. *Are standardized, psychometrically sound (acceptable reliability and validity described in the text) outcome measures used?* (Leber et al., 1996) One point will be awarded when the outcome measure instrumentation involves psychometrically sound instruments that have been validated and shown to be reliable. Basic psychometric properties must be presented in the text of this article to receive this point. (No points will be awarded for instruments that are developed solely for the present study unless the researchers present evidence that the instrument is psychometrically sound.)

13. *Are outcomes measured at multiple levels (attitudinal, behavioral, and observational)?* (Heneghan et al., 1996) One point will be awarded if outcomes are measured across at least two different levels. Therefore, a study measuring outcomes only related to attitudes will not receive a point. For purposes of this study, self-report measures of behaviors (e.g., a violent behavior checklist) will be considered a different level than self-report measures of attitudes. However, self-report measures of behavior will only be counted when they relate to actual behavior (i.e., only behaviors that have actually occurred), and these will not be counted if they measure hypothetical behavior (e.g., a participant is presented with a case study situation and asked how he or

she might respond in that situation). Measures of hypothetical behavior will be regarded as attitudinal measures for purposes of this study.

14. *Are appropriate statistical analyses used?* (Leber et al., 1996) One point shall be awarded for the use of statistical analyses that are sensitive to measure change over time, as in repeated measures ANOVA or regression. Therefore, a study without a control group could receive a point in this category if the statistical analyses measure pretest–posttest change. A study with a control group will only receive a point in this category if the statistical analyses account for both change over time and group comparisons. For example, statistical analyses using paired *t* tests for posttest measures only will not receive a point.

15. *Is a follow-up assessment included (at least one month following the intervention)?* (Guterman, 2004) One point will be awarded for studies that include (in addition to immediate posttest assessment) an additional follow-up assessment at least 1 month following the completion of the intervention.

References

- American Educational Research Association. (2006). *Review of educational research submissions*. Retrieved April 6, 2006, from <http://www.aera.net/publications/?id=505>
- Avery-Leaf, S., Cascardi, M., O'Leary, K. D., & Cano, A. (1997). Efficacy of a dating violence prevention program on attitudes justifying aggression. *Journal of Adolescent Health, 21*, 11-17.
- Budde, S., & Schene, P. (2004). Informal social support interventions and their role in violence prevention: An agenda for future evaluation. *Journal of Interpersonal Violence, 19*, 341-355.
- Centers for Disease Control and Prevention. (2000). *CDC funds ten prevention projects on sexual violence and intimate partner violence among racial and ethnic minority populations*. Retrieved April 6, 2006, from <http://www.cdc.gov/OD/OC/MEDIA/pressrel/r2k1010a.htm>
- Chalmers, T. C., Smith, H., Blackburn, B., Silverman, B., Schroeder, B., Reitman, D., et al. (1981). A method for assessing the quality of a randomized control trial. *Controlled Clinical Trials, 2*, 31-49.
- Christensen, A., & Heavey, C. L. (1999). Interventions for couples. *Annual Review of Psychology, 50*, 165-190.
- Coker, A. L. (2004). Primary prevention of intimate partner violence for women's health: A response to Plichta. *Journal of Interpersonal Violence, 19*, 1324-1334.
- Corcoran, K., & Fischer, J. (2000). *Measures for clinical practice: A sourcebook. Volume 1: Couples, families, and children*. New York: Free Press.
- DeVoe, E. R., & Kantor, G. K. (2002). Measurement issues in child maltreatment and family violence prevention programs. *Trauma, Violence, and Abuse, 3*, 15-39.
- Dickersin, K. (2002). Systematic reviews in epidemiology: Why are we so far behind? *International Journal of Epidemiology, 31*, 6-12.
- Foshee, V. A., Bauman, K. E., Arriaga, X. B., Helms, R. W., Koch, G. G., & Linder, G. F. (1998). An evaluation of Safe Dates, an adolescent dating violence prevention program. *American Journal of Public Health, 88*, 45-50.
- Foshee, V. A., Bauman, K. E., Ennett, S. T., Linder, G. F., Benefield, T., & Suchindran, C. (2004). Assessing the long-term effects of the Safe Dates Program and a booster in preventing and reducing adolescent dating violence victimization and perpetration. *American Journal of Public Health, 94*, 619-624.
- Foshee, V. A., Bauman, K. E., Greene, W. F., Koch, G. G., Linder, G. F., & MacDougall, J. E. (2000). The Safe Dates Program: 1-year follow-up results. *American Journal of Public Health, 90*, 1619-1622.
- Fullwood, P. C. (2002). *Preventing family violence: Community engagement makes the difference*. San Francisco: Family Violence Prevention Fund. Retrieved September 29, 2005, from <http://endabuse.org/programs/children/files/Preventing.pdf>
- Gay, L. R., & Airasian, P. (2000). *Educational research: Competencies for analysis and application* (6th ed.). Upper Saddle River, NJ: Merrill.
- Gallagher, E. J. (1999). Systematic reviews: A logical extension of evidence-based medicine. *Academic Emergency Medicine, 6*, 1255-1260.
- Gil, D. G. (1996). Preventing violence in a structurally violent society: Mission impossible. *American Journal of Orthopsychiatry, 66*, 77-84.

- Guterman, N. B. (2004). Advancing prevention research on child abuse, youth violence, and domestic violence: Emerging strategies and issues. *Journal of Interpersonal Violence, 19*, 299-321.
- Heneghan, A. M., Horwitz, S. M., & Leventhal, J. M. (1996). Evaluating intensive family preservation programs: A methodological review. *Pediatrics, 97*, 535-542.
- Hickman, L. J., Jaycox, L. H., & Aronoff, J. (2004). Dating violence among adolescents: Prevalence, gender distribution, and prevention program effectiveness. *Trauma, Violence, & Abuse, 5*, 123-142.
- Jaffe, P. G., Suderman, M., Reitzel, D., & Killip, S. M. (1992). An evaluation of a secondary school primary prevention program on violence. *Violence and Victims, 7*, 129-146.
- Lavoie, F., Vezina, L., Piche, C., & Boivin, M. (1995). Evaluation of a prevention program for violence in teen dating relationships. *Journal of Interpersonal Violence, 10*, 516-524.
- Lee, R. K., Sanders Thompson, V. L., & Mechanic, M. B. (2002). Intimate partner violence and women of color: A call for innovations. *American Journal of Public Health, 92*, 530-534.
- Leber, D., St. Peters, M., & Markman, H. J. (1996). Program evaluation research: Applications to marital and family therapy. In D. H. Sprenkle & S. M. Moon (Eds.), *Research methods in family therapy* (pp. 485-506). New York: Guilford.
- Lilford, R. J., Richardson, A., Stevens, A., Fitzpatrick, R., Edwards, S., Rock, F., et al. (2001). Issues in methodological research: Perspectives from researchers and commissioners (Executive summary). *Health Technology Assessment, 5*, 1-6.
- Macgowan, M. J. (1997). An evaluation of a dating violence prevention program for middle school students. *Violence and Victims, 12*, 223-235.
- Matthews, N. A. (2000). Generic violence prevention and gendered violence. *Violence Against Women, 6*, 311-331.
- McFarlane, J., Soeken, K., & Wiist, W. (2000). An evaluation of interventions to decrease intimate partner violence to pregnant women. *Public Health Nursing, 17*, 443-451.
- Mitchell-Clark, K., & Autry, A. (2004). *Preventing family violence: Lessons from the Community Engagement Initiative*. San Francisco: Family Violence Prevention Fund. Retrieved September 29, 2005, from <http://endabuse.org/programs/children/files/Preventing2.pdf>
- Nabi, R. L., Southwell, B., & Hornik, R. (2002). Predicting intentions versus predicting behaviors: Domestic violence prevention from a theory of reasoned action perspective. *Health Communication, 14*, 429-449.
- Osofsky, J. D. (2003). Prevalence of children's exposure to domestic violence and child maltreatment: Implications for prevention and treatment. *Clinical Child and Family Psychology Review, 6*, 161-170.
- Plichta, S. B. (2004). Intimate partner violence and physical health consequences. *Journal of Interpersonal Violence, 19*, 1296-1323.
- Ramsay, J., Richardson, J., Carter, Y. H., Davidson, L. L., & Feder, G. (2002). Should health professionals screen women for domestic violence? Systematic review. *British Medical Journal, 325*, 314-326.
- Rynerson, B. C., & Fishel, A. H. (1993). Domestic violence prevention training: Participant characteristics and treatment outcomes. *Journal of Family Violence, 8*, 253-366.
- Sabol, W. J., Coulton, C. J., & Korbin, J. E. (2004). Building community capacity for violence prevention. *Journal of Interpersonal Violence, 19*, 322-340.
- U.S. Preventive Services Task Force. (2004). *Screening for family and intimate partner violence*. Rockville, MD: U.S. Department of Health and Human Services. Retrieved May 27, 2007, from <http://www.ahrq.gov/downloads/pub/prevent/pdfser/famviolser.pdf>.
- Wallace, H. (2005). *Family violence: Legal, medical, and social perspectives* (4th ed.). Boston: Allyn & Bacon.
- Weisz, A. N., & Black, B. M. (2001). Evaluating a sexual assault and dating violence prevention program for urban youths. *Social Work Research, 25*, 89-100.
- Wolfe, D. A., & Jaffe, P. G. (1999). Emerging strategies in the prevention of domestic violence. *Future of Children, 9*, 133-144.