

Evaluating batterer counseling programs: A difficult task showing some effects and implications

1.1 Introduction

1.1.1 Introduction

The question of what to do about men who batter their female partners has haunted the domestic violence field since its emergence in the late 1970s. Many advocates working with battered women felt—and many still feel—that few batterers could be changed given the social reinforcement for and tolerance of violence against women (Taubman, 1986). Trying to counsel or educate such men might, in fact, raise false hopes in battered women and worsen their already difficult circumstances. Protection for women and separation from their male batterers, therefore, became the overarching intervention objective (e.g., Dobash & Dobash, 1992). In the late 1970s, however, some men's counselors allied with the battered women's advocates and began antisexist consciousness-raising groups primarily for men who professed wanting to change (Adams, 1988). Group facilitators led discussions that exposed men's socialization to dominate women and, in some cases, use violence to maintain that dominance (see Adams, 1989).

Batterer programs gradually became more sophisticated by adopting cognitive-behavioral techniques from the counseling of other violent men (Pence & Paymar, 1993). Some of these programs emphasizing gender issues have been accused of being too "confrontive" (Stosny, 1995), while others emphasizing skill building are often criticized for being naïvely superficial and lacking a clear message of change (Gondolf & Russell, 1986). The fact is that numerous curriculums have been developed incorporating both gender issues and cognitive-behavioral techniques (e.g., Kivel, 1992; Pence & Paymar, 1993; Russell, 1995; Stordeur & Stille, 1989).

While a gender-based cognitive-behavioral orientation is the most prominent approach, other more psychodynamic or emotive approaches have been forwarded as well (e.g., Dutton, 1998; Stosny, 1995). These approaches attempt to address the psychological issues and emotional hurts of men that may contribute to their abuse. A variety of other therapies have also emerged: anger management (a streamlined others.2Pence & m10.1((a)-410v538oheie -

these sorts of issues are likely to further their validity, and those that at least acknowledge them will help clarify interpretation of the results.

The first major set of issues is those that deal with program definition. These issues are probably the least discussed and refined in the evaluations of batterer programs thus far. Is the “program” merely the group counseling within a weekly, 1.5- to 2-h session? Is it also the program intake, individual assessment or screening, dismissal policies, and reporting procedures? Does it include the outreach to and support of the victims that some programs offer? The evolving nature of especially community-based programs complicates the answer to these questions. Protocols, curriculums, counselor skill, administration, and referral and reporting procedures tend to change and develop. Does the evaluation outcome at the end of an extended follow-up apply to the current program that may have evolved from the initial one? Our multisite study documented substantial changes at the program sites following our period of subject recruitment and treatment (see [Gondolf, 2002a](#)). New program directors are currently at three of the four sites, half to all of the staff have been replaced, one program has an entirely new curriculum, the two 3-month programs have been extended to 4 and 6 months, and the court referral system at two of the sites has been substantially revised.

Additionally, batterer programs are enmeshed in an elaborate intervention system that includes police practices, court action, probation supervision, civil protection orders, victim services, additional services for the men, community resources, and local norms (see [Gamache, Edleson, & Schock, 1988](#); [Murphy, Musser, & Maton, 1998](#)). These components ideally work together within what has come to be called a “community coordinated response” ([Pence & Shepard, 1999](#)). This sort of system varies considerably in extent, development, and actual coordination under so-called domestic violence councils. The councils also run the gambut in participation, viability, and authority. To what extent do we separate and distinguish batterer counseling from these components, and to what degree do the components contribute to the counseling? Some observers, in fact, argue that such

The outcome period is also a crucial issue. What is considered a sufficient duration to test—6 months, 1 year, 2 years? When does the period start—at program intake or after the program when the men have received a sufficient “dose” of counseling? If we have programs of different lengths, should we adjust the outcome periods so that they have equivalent postprogram follow-ups? One of the answers emerging from the evaluations in other fields is the shift from cumulative outcomes to longitudinal retrospective ones (i.e., how many months has the program participant been violent-free *p i o* to the follow-up endpoint). For example, “recovery” from alcohol misuse appears to be nonlinear with points of relapse. The current trend in alcohol treatment evaluation is, therefore, to measure periods or days of sobriety or current alcohol use rather than the cumulative rate of relapse (e.g., [National Opinion Research Center, 1996](#)). These different approaches to measuring outcome can shift apparent “failure” to “success,” as we will illustrate later with our multisite evaluation.

Finally, we are faced with a fundamental issue of interpretation: How much of an effect constitutes success or warrants an endorsement (see [Jennings, 1990](#))? What do we mean by “effective” or that a program “works.” Program evaluators have questioned: Effective compared with what, with whom, and under what circumstances? There is the additional question being raised in these days of “managed care:” Is the effect worth the cost to insurance companies or state agencies? In the case of batterer programs: Is the cost diverting needed and deserved funding from victim services, or is it helping the victims through beneficial outcomes? The answers to these questions are largely subjective and require an interpretation of the outcome using the nuance of qualitative research and clinical experience, as well as a cost analysis (see [Jones, 2000](#)).

2. . . . *i*

2. .1. *E p i t i*

The biggest and most controversial issue surrounding program evaluation is whether to use experimental or quasi-experimental designs. While the former is considered more “scientific,” experimental designs encounter substantial implementation challenges and conceptual shortfalls. Some evaluators question whether they can truly be achieved in criminal justice settings and end up causing more disruption and misinformation than they are worth ([Marshall & Serran, 2000](#)). But can quasi-experimental and other alternative designs meaningfully address program effectiveness without an equivalent control group? As illustrated below, our experience is that current analytical procedures make it possible to simulate experimental conditions within quasi-experimental evaluations and to do so in a way that represents a “dose response” to counseling.

Most researchers in the positivist tradition hold experimental designs to be the “gold standard” for evaluation ([Boruch, Snyder, & DeMoya, 2000](#); [Dunford, 2000a](#)). Randomly assigning subjects to an experimental (i.e., treatment) group and a control (i.e., nontreatment) group, as a means to maximize internal validity, seems straightforward enough. Unfortunately, implementation of such designs is especially problematic, as an extensive literature in the medical and public health field shows. But experimental evaluations raise conceptual issues, as well. The primary one has to do with the “intention to treat” approach assumed in

most experimental designs, as opposed to a dose response (Efron & Feldman, 1991). In experimental designs, the experimental group consists of everyone sent to the program or treatment, whether they receive the treatment or dropout. The drop outs—or those with a “low dose” of the program—can cancel out the apparent effectiveness of the program completers—or “high dose” participants. The comparison of an experimental group versus control group, therefore, may tell less about treatment effectiveness and more about the procedures of referring to and retaining men in a certain program.

The effectiveness of the referral may be as much dependent on the intervention system as a whole as it is on the program. Some programs screen out or dismiss men who do not appear “motivated” to change; others attempt to motivate men in program orientations or through the threat of sanctions. However, most batterer programs monitor the men’s attendance and return them to the court system if they do not attend. These programs contend that the responsibility to keep the men in the program is not theirs but rather the responsibility of the courts that refer the men. There are analytical steps to assess the dose response in an experimental design, but they essentially turn the design into a quasi-experiment and require large samples and extensive intake assessment to accomplish.

Another major implementation issue, especially in the courts, is finding a pure control group. Men assigned to open probation are generally used as the control group, but many of these men, or their partners, receive compensating treatments, sanctions, or surveillance in a viable “coordinated community response.” For instance, we found women whose partners dropped out early were more likely to seek additional help from a women’s center. Moreover, an experiment can be disruptive to the intervention system and actually change it. The results may end up reflecting an altered system rather than the initially intended one. This issue has especially been evident in recent experimental evaluations of batterer programs (e.g. kre16(programs)-2

control variables and a large sample of respondents. But a more fundamental problem leads statisticians to consider logistic regression as a “naïve” analysis. The regression treats program dropout or attendance as an independent variable when it is “endogenous” or dependent on many of the same controls that influence reassault. Statisticians in the public health field have long faced this problem since so many community-based interventions do not readily lend themselves to experimental evaluation (e.g., how would you test to see if preventing teen pregnancy improves adult physical health). One popular option has been the use of Propensity Score analysis which develops a weighting for the likelihood to drop out and uses that weighting, or “score,” to match subjects in the dropout and completer groups (Jones, D’Agostino, Gondolf, & Heckert, in press). (A variation of Propensity Score analysis directly adjusts the effect of program completion on reassault.)

2. . . i i i

An alternative view is that both pure and quasi-experiments are not sufficiently “realistic” (Dobash & Dobash, 2000; Pawson & Tilly, 1997; Van Voohis, Cullen, & Applegate, 1995). That is, they often manipulate the intervention for random assignment or impose a prescribed treatment. But most of all, they fail to account for the context of the program that may substantially contribute to the programs outcomes. This context includes not only the immediate context of the intervention system components but also additional interventions, help-seeking, and circumstances (e.g., drinking, unemployment) during the follow-up. A more realistic outcome might be, therefore, a dynamic conception that includes mediating individual conditional factors and community resources and supports (see Mulvey & Lidz, 1985; Steadman, 1982). The outcomes may be influenced, for example, by women’s counseling or men’s alcohol treatment during the follow-up or by the police practices and service availability in the community.

Computer modeling can be done to account for the dynamic or time-varying nature of the follow-up. Longitudinal data, of course, are needed not only of the outcome but also of conditional or mediating variables, along with predisposing characteristics that are both static (e.g., race, personality) and dynamic (e.g., employment, drinking). One popular analytic tool for this purpose is Generalized Estimating Equations (GEE; Laing & Zeger, 1986). It derives an equation for the pooled follow-up intervals, and the correlation between those intervals, using a series of equations predicting the outcome of each separate follow-up interval outcome (thus producing consistent standard errors). GEE also accommodates so-called censored data of individuals who do not complete the full follow-up.

The more complex issue is how to account for the broader context, while controlling for the differences in program dropouts and completers. One approach, also developed in the public health field, is Instrumental Variable (IV) analysis (e.g., Bollen, Guilkey, & Mroz, 1995). In this structural equation approach, predisposing characteristics are entered in an equation to predict dropout along with a set of contextual “instrumental variables” that uniquely distinguish dropout (e.g., victim services, referral source, and perceived sanctions). This “dropout” equation is then entered into a second equation predicting reassault, along with predisposing variables and IVs unique to reassault (e.g., additional services, unemployment, and police domestic violence arrests). This approach also accounts for potential

24 . i₀ i₀ t t - i

24 .1. - t - t

Over 40 batterer program evaluations have been published in academic journals, and many more are available as agency reports. Reviews of the batterer program evaluations show 50–80% of the program completers to be nonviolent at the end of a 6-month to 1-year period, according to their female partners (e.g., Edleson, 1996; Gondolf, 1997a; Rosenfeld, 1992; Tolman & Bennett, 1990). The reduction of other forms of abuse is less clear, but one study showed that only about 40–50% of the participants had stopped their terroristic threats at a 6-month follow-up (Edleson & Syers, 1990). It may be that some men displace their physical abuse to heightened verbal and psychological abuse. The batterer program “success rates,” while compromised by high dropouts and incomplete follow-ups, are comparable to those in drunk driving, drug and alcohol, sex offender, and check forging programs (see Furby, Blackshaw, & Weinrott, 1989; Hubbard et al., 1989; Kassebaum & Ward, 1991; Schare, 1992). In a social science “court,” most all these batterer program evaluations would, however, be dismissed on technicalities or as circumstantial evidence (see Quinsey, Harris, Rice, & Lalumiere, 1993). They are compromised by selection bias, low response rates, short follow-up periods, no or weak control groups, and no calculations of effect size.

The effectiveness of batterer programs has been addressed in a more comprehensive way using meta-analyses of both experimental and quasi-experimental evaluations (Durlak & Lipsey, 1991).² These analyses statistically summarize the magnitude of the effect attributable to the group of programs considered in the analyses. Cohen (1977) developed statistical calculations for effect size as an alternative to the dependence on and misuse of t values. In the case of batterer programs, the coefficient for “effect size”—most commonly Cohen’s —

pose several caveats (Wilson & Lipsey, 2001). The effect size is not, therefore, a “bottom line” of program success or failure independent of interpretation, although it is sometimes used that way in these analyses.

The most current meta-analysis selected 15 of the most methodologically sound batterer program evaluations and included the three recent experimental evaluations discussed below. It found, similar to the other previous meta-analyses, only a small effect size for batterer programs overall (Babcock & Robie, in press). Effect sizes (computed with the d statistic) were .10 for the experimental designs ($n = 8$ effects), .41 for the quasi-experimental

2.1.1. Experimental Limitations

The methodological limitations of the experimental designs may be a substantial reason for the slight effect size found in the meta-analysis. As a special issue of *Criminal Justice and Behavior*, Dunford (2000) argues, the experimental evaluation of batterer programs should get the most weight, since their designs are the most rigorously scientific (Boruch et al., 2000; Dunford, 2000a). A careful look at the three recent experimental evaluations exposes, however, implementation problems that compromise their results (for details, see Gondolf, 2001). Both the New York City (Davis, Taylor, & Maxwell, 1998) and South Florida (Feder & Dugan, 2002) evaluations had substantial problems implementing the random assignment. As much as 30% of the potential cases were overridden at one site, and the point of random assignment was moved twice at the other. Moreover, legal opposition to random assignment disrupted the South Florida evaluation (Feder et al., 2000). The batterer program in New York was changed during the subject recruitment.

There were also major problems in implementing the follow-up. The response rate of female partners was only 20% in the South Florida evaluation, which led to the use of confounded probation records as the outcome source. The New York City evaluation used private investigators to track down some nonrespondents, which may have affected their disclosure. The low program completion—as low as 40% for the South Florida program—exacerbated the intention-to-treat distortions of the experiments, as well, and raises questions about the program operation in general. (Completion rates at programs in our multisite evaluation were 55–70% varying with the required length of the program (Gondolf, 1997b).)

A “clinical trial” experiment at a Navy base showed no effects for batterer counseling compared with other options such as safety planning and intensive monitoring (Dunford, 2000b). This evaluation is difficult to generalize to the civilian public, however, because of the different characteristics of the Navy men, and the heavy and certain sanctions for reassault while under Navy supervision. Moreover, the comparative options in the Navy experiment, such as the monitoring and safety planning, are incorporated as part of the intervention system in most civilian batterer programs rather than being alternatives. Consequently, the results of the previous experiments, although instructive, must be viewed and applied with caution.

In summary, it is difficult to make a categorical recommendation about batterer program effect given the methodological problems of many of the evaluations and the limitations of effect-size statistics. The meta-analyses do, however, raise suspicions about excessive claims of success; we have no substantial evidence that most programs are effective or that any programs are highly effective.

3. A multisite evaluation of batterer intervention

3.1. Introduction

We developed a multisite evaluation in an effort to address some of the conceptual issues and methodological shortcomings of previous evaluations (Gondolf, 1997b, 1999a, 2000d,

2001). This evaluation had the advantage of substantial funding from the Center for Disease Control and benefit of both clinical and research advisors. In our opinion, the results of the evaluation indicate that at least some programs are effective in stopping assault and abuse and that batterer intervention in general shows some promise. The evaluation also suggests that the predominant gender-based cognitive-behavioral approach to counseling may be appropriate for the majority of men. The slight effects shown in other evaluations may be the result of poorly implemented programs or intervention systems.

Our multisite evaluation employed a naturalistic comparative design across sites, with quasi-experimental studies within the sites. We identified programs in four cities that represented a range in program format and system components. The four interventions were based on “well-established” batterer programs that used the prevailing gender-based, cognitive-behavioral approach. The intervention systems ranged, however, from a streamlined system of (1) pretrial referral to batterer counseling at a preliminary hearing, (2) additional referrals for court-identified alcohol or psychological problems, (3) 3 months of instructional group counseling, and (4) periodic court review of program compliance. At the other extreme was a comprehensive system based on (1) postconviction referral, (2) individual assessment, (3) in-house alcohol and psychological treatment, (4) 9 months of discussion-oriented group counseling, and (5) monitoring of compliance by probation officers. The remaining two systems included a 3-month discussion-oriented program and a 5 1/2-month instructional program, both with postconviction referrals.

The evaluation consisted of a 4-year follow-up, starting at program intake, with periodic interviews of 840 men and their female partners. During 1995, research assistants recruited program participants from the four sites ($n = 840$) and administered a uniform set of background questionnaire, personality inventory (MCMI-III; Millon, 1994), and alcohol test (MAST; Selzer, 1971). The sample size was reduced in the extended follow-up (15–48 months) to 618. The participants who were voluntary and recruited in the first 2 months were deleted. This deletion furthered the focus on the court-referred men and made the follow-up less costly. (The reassault rates presented here are those computed for the court-referred men alone.) A variety of qualitative and quantitative outcomes, and conditional or time-varying variables, was assessed at 3-month intervals for 4 years. The response rate with the female partners of the program participants was approximately 70% for the first 30 months and 60% for the full 4 years (for human subjects and tracking procedures, see Gondolf, 2000b).³

The principal outcome of reassault was based on reports from the men’s initial new female partners, and confirmed with an analysis of police reports and men’s self-reports (Heckert & Gondolf, 2000a,b), a coded qualitative analysis of women’s narratives (Heckert,

³ The most persistent concerns about outcomes in batterer program evaluations are the reliability of women’s reports. Don’t the women deny or withhold the truth of being reassaulted? We believe that disclosure was facilitated by our frequent and long-term contact, the funnel questioning that allowed women to tell their “story,” and debriefing interviews that inquired about obstacles or barriers. All of the women had, moreover, already officially reported (or had someone report for them) their initial assaults and had to appear in court to testify about

Matula, & Gondolf, 2000), and a Capture–Recapture analysis of self-reports and arrest records (Gondolf, Chang, & LaPorte, 1999). Several attrition analyses, including one using Heckman (1979) regressions, revealed no significant response bias (Jones, 1998). The scope of the data, size of the pooled sample, and longitudinal follow-up enabled us to conduct some more complex analyses of the dynamic outcome (Jones & Gondolf, 2001) and the contextualized program effect (Gondolf & Jones, 2001). We also monitored the counseling sessions and conducted periodic observations and interviews about the intervention system and community context (see Gondolf, 2002a, chap. 4).

.2. / /

lower rearrests rates for program completers (Babcock & Steiner, 1999), as does our comparison of arrest rates for different court actions at one of our sites (Gondolf, 1998). At the minimum, the court referral offers an intensified supervision through the program and a weekly reprieve for the men's partners.

... o III¹

The second and more debated question is whether the counseling is effective over no counseling—that is, is there a program or treatment effect beyond the arrest and court appearance? When we compared those men receiving a minimum dose of counseling (less than 2 months) with those who completed 2 months or more, we found a 50% reduction in reassault for the “completers” (36% of the completers reassaulted vs. 55% of the dropouts). This reduction amounts to an effect size of .48 (Babcock & Robie, *in press*). The site with the streamlined system was deleted from this analysis because the quasi-control group was contaminated by the court-review process. Most dropouts at this particular site received compensating interventions that contained their potential reassault (for documentation, see Gondolf & Jones, 2001). The reduction in reassault is even higher when controlling for the men living with their partners (66% reduction: 40% vs. 67%). Many men drop out because of no longer being with their partner and not seeing the relevance of counseling.

Different characteristics between the dropouts and completers may, of course, account for the apparent reduction attributed to the program. When controlling for demographic, relationship, personality, and behavioral characteristics, in a logistic regression (or naïve analysis), program completion remains significant with a reduction of 18% in reassault.⁴ To account for the endogenous nature of dropout, we also conducted the form of structural equation analysis called IV analysis. According to this approach, program completion produced an effect size of .44–.64, depending on the model specification (Gondolf & Jones, 2001)

interviews and observations of batterer programs across the country promoted a comprehensive approach (Gondolf, 1985). According to feminist and social learning theories, men are, furthermore, socialized by a male-dominated society to exert power over women and reinforced by a violent culture (e.g., Taubman, 1986). It would take a long and concentrated effort to “unlearn” this socialization. However, the reassault and related outcomes across the four sites were strikingly similar (Gondolf, 1999a). When we controlled for demographics, relationship status, personality, and previous behavior, we still found no significant site effect on reassault.

The lack of a site effect was confirmed in the IV analysis which tested for a program effect (Gondolf & Jones, 2001), and in the GEE analysis identifying predictors of reassault (Jones & Gondolf, 2001). At the 15-month follow-up, there was a slight decrease in reassault in the longer programs, but this was not corroborated by other indicators and disappeared at the cumulative 30-month follow-up (Gondolf, 1999a; Gondolf, 2000d).⁵ A similar result regarding the effectiveness of program length has appeared in evaluations of sex offender treatments. Longer treatment programs have not been shown to be more effective than shorter treatment programs (Marshall & Serran, 2000).

There are several possible explanations for the lack of a “site effect.” One explanation reflects the assumptions of brief therapy and managed care: The men who are going to change begin to do so within 3 months of counseling, and those who might benefit from longer counseling tend to drop out. Indeed, 3 months represented not only the duration of two of the programs, but also a threshold of dropout for the longer programs (Gondolf, 1997b). Another explanation is the possible uniqueness of each intervention system. The NIMH multisite depression study found an interaction between site and counseling approach that appeared to be related to cultural and resource differences among the sites (Elkin et al., 1989). The batterer intervention systems in our evaluation similarly evolved in response to a distinctive set of circumstances, resources, expectations, and personalities. For example, the site of the longest and process-oriented program had more therapists per capita than all but one other city in the country. This circumstance may account for the batterer program participants at this site being the most likely to have previously received counseling and expect longer, process-oriented batterer counseling.

The most compelling explanation for no site effect, however, may be the compensating influence of system components. The streamlined system operated much like the increasingly popular drug courts. Its court supervision of the cases ensured a swift and certain response to noncompliance (Gondolf, 2000c). Under the pretrial referral, the men entered the program in an average of 2 1/2 weeks after arrest, as opposed to several months at the postconviction systems, and they had to reappear in court periodically to confirm their program attendance. This system dramatically reduced no-shows (from 30% to 5%) and sustained a high completion rate of 70% despite the coerced attendance. The system appears to matter.

⁵ The distribution of psychological problems and previous behavior was strikingly similar across the four sites, but socioeconomic status was lowest for the streamline site and greatest for the comprehensive site (see Gondolf, 1999b).

3. *Program Intake*

A secondary finding points to areas for improvement. Approximately a quarter of the men (about half those who reassaulted during the four-year follow-up) reassaulted their partners more than once (Gondolf, 2000d). Most of these men began their reassaults shortly after program intake, and were responsible for over 80% of the injuries. They were, not surprisingly, more likely to have been severely violent in the past and to have been arrested or treated previously. However, their personality profiles did not distinguish them from men who reassaulted only one time, and men who did not reassault (Gondolf & White, 2001). Their patterns of violence, based on coding of the women's narratives, also did not distinguish these particularly dangerous men (Gondolf & Beeman, 2003). The analysis of women's narratives did, however, reveal distinct patterns and trajectories of violence that warrant further consideration. Calibrated identification of violence patterns might improve the prediction of further violence. This possibility makes sense given that the gross indicators of antisocial behavior continue to be correlated with program outcome—"past violence is the best predictor of future violence" (e.g., Dutton, Bodnarchuk, Kropp, Hart, & Ogloff, 1997).

In summary, the repeat reassaulters did not appear as a distinct "batterer type." The most distinguishing factor was the lack of response to these men. Their partners were less likely to take action, possibly out of fear or subjection; and further arrests, protection orders, and treatment were not as likely, as a laboratory study found with more antisocial men (Jacobson & Gottman, 1998). More extensive case management and systematic victim contact might help expose repeated reassaults, and certain and decisive intervention for a initial reassault might make a difference on the outcome. It would likely reduce repeated reassaults.

Another option is to attempt to identify these "high-risk" men at program intake, as several risk assessment instruments attempt to do (for overviews of these instruments, see Dutton & Kropp, 2000; Roehl & Guertin, 2000). Risk assessment instruments have been shown to improve prediction over clinical judgment, but their predictive power still falls far short of what could be considered reliable. We attempted to identify predictors for reassault and especially for repeat reassault. Static intake variables showed little predictive power for reassault during the crucial first 15 months (Jones & Gondolf, 2001). The only significant intake predictors were severe psychological problems, previous severe abuse, as other research of violent men has found (Walters, 2000). The only substantial predictor was drunkenness during the follow-up. (Those men identified by their partners as "drunk," at least once during a follow-up interval, were four times more likely to reassault.) We also found at least a weak association between accompanying alcohol treatment and a reduction in reassault; however, the treatment varied considerably in extent and intensity and was not sustained with many of the most violent men.

This latter finding lends support to the recurrent association of alcohol misuse to woman assault. The relationship between alcohol and assault is, however, a complex one, as numerous reviews of the topic have indicated (e.g., Gondolf, 1995). We were able to demonstrate the separation between "drunkenness" and reassault by lagging the drunk-

enness in the analysis and also with a separate variable for “drinking prior to the assaultive incident.” The drunkenness, therefore, appears to be more of a matter of lifestyle, rather than a direct cause. Interestingly, other research has shown that the association of drinking to reassault disappears when male attitudes of dominance are controlled (Johnson, 2000). The heavy drinking and reassault may both be symptoms of the same underlying attitudes or lifestyle. Some researchers have, more broadly, implicated unstable lifestyles as predictive of reassault by violent offenders in general (Hanson & Wallace-Capretta, 2000).

.. i h o p i h o t i

intervention. This has been done in NIAAA's elaborate test of specialized treatment for different types of alcoholics (Project Match, 1997). This multisite study concluded that "patient-treatment matching . . . adds little to enhance outcome treatment" (Gordis, 1997, p. 6).

To explore the utility of batterer types, we created personality-based types in two ways: (1) with a cluster analysis of MCMI-III subscales, and (2) with a classification of the MCMI-III profiles based on interpretative guides (White & Gondolf, 2000). Both approaches produced "types" that reflected tendencies suggested in the prevailing typology (Holtzworth-Munroe & Stuart, 1994). The most remarkable part was the percentage of men who showed no "psychopathology" at all, and the diversity of profiles overall, as found with clustering using the MCMI-II (Hamberger, Lohr, Bonge, & Tolin, 1996). We did not find any evidence to substantiate a prevailing "abusive personality" with underlying tendencies of a borderline personality (Dutton, 1998).⁶

Our types are not, however, as complex as the personality/behavioral types in the most current typology research (Holtzworth-Munroe, Meehan, Herron, Rehman, & Stuart, 2001). The violent and criminal behavior incorporated in the more complex batterer types may make the types appear predictive, as has been suggested with measures of psychopathy (Skeem & Mulvey, 2001). Antisocial behavioral indicators (e.g., prior violence and

Our findings shift the focus from curriculum change or treatment diversification toward program structure and system coordination. In terms of program structure, program intensity rather than extent warrants more attention. More might be done to monitor and contain men during the first few months after program intake, when men are most likely to first reassault. Much like the move toward intensive outpatient treatment in the alcohol field, some men might be required to attend three to four long sessions per week, rather than the usual once-a-week sessions of 1.5–2 h. The men who have reassaulted, or have severely assaulted their partners in the past, are logical candidates for these sessions. Programs might also offer more continuous support to the men's female partners through women's advocates or caseworkers. The outreach to the women could be substantially improved within the intervention systems of our study. While about a third of the batterers' female partners reported some contact with battered women's services within the first 3 months of program intake (beyond contact with a legal advocate), only 8% of the women had any contact in the next 12 months (Gondolf, 2002b). Most of this later contact was "reactive" in that it was a response to additional assaults. Consequently, women services were not associated with a reduction in reassault (Jones & Gondolf, 2001). Additional support might develop ongoing risk management that would bring additional intervention and assistance as appropriate, especially in cases of repeat reassault. It might also help women more assertively access and benefit from available resources and services.

Several areas need attention within the intervention system as a whole. More men need to be moved swiftly into batterer programs. The effectiveness of programs is apparently

detrimental to the drinker as well as to others. Some different counseling approaches or formats may also be needed for nonarrested men. A subsample of “voluntary” participants in our evaluation tended to drop out of conventional batterer treatment and reassault their partners (Gondolf, 1997b). They were also more likely to have severe psychological problems, particularly depression.

An additional consideration is the possible need for specialized counseling for African American men and Latino men (see Carrillo & Tello, 1998; Williams, 1994). Our review of the research shows cultural issues that may account for higher dropout and poorer outcomes among these men (Gondolf & Williams, 2001). However, there is little research that documents poorer outcomes in conventional programs or improved outcomes in culturally specific programs. One preliminary study of specialized batterer counseling shows African American men having more favorable impressions of culturally specific counseling ($n = 41$; Williams, 1995). However, the African American men in our multisite evaluation were no more likely to reassault than their white counterparts. To examine this issue more systematically, we have begun a clinical trial of specialized counseling for African American men arrested for domestic violence. This evaluation, funded by the National Institute of Justice, compares conventional batterer

At a minimum, evaluation might consider the influence of system components on outcomes. We do not know, for instance, the effectiveness of a comprehensive, 9-month program, coupled with the swift and certain response of a “drug-court type” referral and supervision. We do not know the trade-offs between pretrial referral versus fast-track prosecution, or between court review versus intensified probation. We do not know if these options are unique to certain court jurisdictions and community settings. The National Institute of Justice has launched a multisite demonstration project of more highly coordinated, and even integrated, intervention systems. The evaluation component attached to this project may offer some answers in this regard.

Finally, there remains the larger question of how best to develop counseling programs and intervention systems. The administration of programs varies tremendously from a clinician-led set of groups to agency-based counseling sessions. Some communities have a scattering of programs competing against each other, and some have one or two programs that maintain a monopoly over the counseling. Some programs have established court liaisons and case managers to help coordinate the intervention components; others rely on specialized probation departments to take this responsibility. It is clear, from our observations, that different administration has different effects on the system and ultimately on the program outcomes (see [Gondolf, 2002a](#)). It is something that also warrants more attention and development. To make some of these most basic developments requires a means to do so.

References

- Adams, D. (1989). Feminist-based interventions for battering men. In L. Caesar, & K. Hamberger (Eds.), *Battering men: A national study of men who batter* (pp. 3–23). New York: Springer.
- Aldarondo, E. (2002). Evaluating the efficacy of interventions with men who batter. In E. Aldarondo, & F. Mederos (Eds.), *Men who batter: A national study of men who batter* (pp. 3-1–3-20). New York: Civic Research Institute.
- Austin, J., & Dankwort, J. (1999). Standards for batterer programs. *Journal of Interpersonal Violence, 14*, 152–169.
- Babcock, J., & LaTaillade, J. (2000). Evaluating interventions for men who batter. In J. Vincent, & E. Jouriles (Eds.), *Men who batter: A national study of men who batter* (pp. 37–77). Philadelphia: Jessica Kingsley Publishers.
- Babcock, J., & Robie, C. (in press). Does batterers’ treatment work? A meta-analytic review of domestic violence treatment. *Criminal Justice and Behavior*.
- Babcock, J., & Steiner, R. (1999). The relationship between treatment, incarceration, and recidivism of battering: A program evaluation of Seattle’s coordinated community response to domestic violence. *Journal of Interpersonal Violence, 14*, 46–59.
- Baumeister, R., Smart, L., & Boden, J. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Journal of Personality and Social Psychology, 70*, 5–33.
- Belenko, S. (1998). Research on drug courts: A critical review. *Criminal Justice and Behavior, 15*, 1–42.
- Bennett, L. (1998). In defense of batterer-program standards. *Journal of Interpersonal Violence, 13*, 93–97.
- Bollen, K., Guilkey, D., & Mroz, T. (1995). Binary outcomes and endogenous explanatory variables: Tests and

- Gamache, D., Edleson, J., & Schock, M. (1988). Coordinated police, judicial and social service response to woman battering: A multi-baseline evaluation across three communities. In G. Hotaling, D. Finkelhor, J. Kirkpatrick, & M. Straus (Eds.), *Coordinated police response to domestic violence* (pp. 193-209). Newbury Park, CA: Sage Publications.
- Gendreau, P., Little, T., & Goggin, C. (1996). A Meta-analysis of the predictors of adult offender recidivism: What works! *Criminal Justice and Behavior*, 13, 575-607.
- Gladwell, M. (2000). *The Tipping Point*. New York: Little, Brown.
- Gondolf, E. (1985). *Men Who Batter: A Comprehensive New Approach to Understanding, Treating, and Preventing Domestic Violence*. Holmes Beach, FL: Learning Publications.
- Gondolf, E. (1987). Changing men who batter: A developmental model of integrated interventions. *Journal of Interpersonal Violence*, 2, 345-369.
- Gondolf, E. (1988). Who are these guys? Toward a behavioral typology of batterers. *Journal of Interpersonal Violence*, 3, 187-203.
- Gondolf, E. (1995). Alcohol abuse, wife assault, and power needs. *Journal of Interpersonal Violence*, 10, 275-284.
- Gondolf, E. (1997a). Expanding batterer program evaluations. In G. Kaufman, & J. Jasinski (Eds.), *Domestic Violence: A National Agenda* (pp. 208-218). Thousand Oaks, CA: Sage Publications.
- Gondolf, E. (1997b). Patterns of reassault in batterer programs. *Journal of Interpersonal Violence*, 12, 373-387.
- Gondolf, E. (1998). *Domestic Violence: A National Agenda*. Report submitted to Pennsylvania Commission on Crime and Delinquency, Harrisburg, PA.
- Gondolf, E. (1999a). A comparison of reassault rates in four batterer programs: Do court referral, program length and services matter? *Journal of Interpersonal Violence*, 14, 41-61.
- Gondolf, E. (1999b). Characteristics of court-mandated batterers in four cities. *Journal of Interpersonal Violence*, 14, 1277-1293.
- Gondolf, E. (1999c). MCMI results for batterer program participants in four cities: Less "pathological" than expected. *Journal of Interpersonal Violence*, 14, 1-17.
- Gondolf, E. (2000a). How batterer program participants avoid reassault. *Journal of Interpersonal Violence*, 15, 1204-1222.
- Gondolf, E. (2000b). Human subject issues in batterer program evaluation. *Journal of Interpersonal Violence*, 15(1), 273-297.
- Gondolf, E. (2000c). Mandatory court review and batterer program compliance. *Journal of Interpersonal Violence*, 15, 437-438.
- Gondolf, E. (2000d). Reassault at 30-months after batterer program intake. *Journal of Interpersonal Violence*, 15, 111-128.
- Gondolf, E. (2001). Limitation of experimental evaluations of batterer programs. *Journal of Interpersonal Violence*, 16, 79-88.
- Gondolf, E. (2002a). *Men Who Batter: A Comprehensive New Approach to Understanding, Treating, and Preventing Domestic Violence*. Thousand Oaks, CA: Sage Publications.
- Gondolf, E. (2002b). Service barriers for battered women with male partners in batterer programs. *Journal of Interpersonal Violence*, 17, 217-227.
- Gondolf, E. (in press). *Men Who Batter: A Comprehensive New Approach to Understanding, Treating, and Preventing Domestic Violence*. Social Work Research.
- Gondolf, E., & Beeman, A. (2003). Are men who repeatedly reassault distinguished by a special kind of violence? *Journal of Interpersonal Violence*, 18, 1-24.
- Gondolf, E., Chang, Y., & Laporte, R. (1999). Capture-recapture analysis of batterer reassaults: An epidemiological innovation for batterer program evaluation. *Journal of Interpersonal Violence*, 14, 191-202.
- Gondolf, E., Heckert, A., & Kimmel, C. (2002). Non-physical abuse among batterer program participants. *Journal of Interpersonal Violence*, 17, 293-315.
- Gondolf, E., & Jones, A. (2001). The program effect of batterer programs in three cities. *Journal of Interpersonal Violence*, 16, 693-704.

- Gondolf, E., & Russell, D. (1986). The case against anger control treatment programs for batterers. *Journal of Interpersonal Violence*, 1(3), 2-5.
- Gondolf, E., & White, R. (2000). "Consumer" recommendations for batterer programs. *Journal of Interpersonal Violence*, 15, 196-215.
- Gondolf, E., & White, R. (2001). Batterer program participants who repeatedly reassault: Psychopathic tendencies and other disorders. *Journal of Interpersonal Violence*, 16, 361-380.
- Gondolf, E., & Williams, O. (2001). Culturally-focused batterer counseling for African-American men. *Journal of Interpersonal Violence*, 16, 283-295.
- Gondolf, E., Yllo, K., & Campbell, J. (1997). Collaboration between researchers and advocates. In G. Kaufman, & J. Jasinski (Eds.), *Collaboration: A new paradigm for social work practice* (pp. 255-261). Thousand Oaks, CA: Sage Publications.
- Gordis, E. (1997, April). Patient-treatment matching—A commentary. *Journal of Interpersonal Violence*, 12, 3.
- Gottfredson, D., Najaka, S., & Kearley, B. (2003). Effectiveness of drug treatment courts: Evidence from a randomized trial. *Journal of Interpersonal Violence*, 18, 171-196.
- Guba, E., & Lincoln, Y. (1989). *Handbook of qualitative research*. Thousand Oaks, CA: Sage Publications.
- Hamberger, L., Lohr, J., Bonge, D., & Tolin, D. (1996). A large sample empirical typology of male spouse abusers and its relationship to dimensions of abuse. *Journal of Interpersonal Violence*, 11, 277-292.
- Hanson, R. K., & Wallace-Capretta, S. (2000). *Domestic violence: A national curriculum*. User Report 2000-06. Ottawa: Department of the Solicitor General of Canada.
- Heckert, A., & Gondolf, E. (2000a). Assessing assault self-reports by batterer program participants and their partners. *Journal of Interpersonal Violence*, 15, 181-197.
- Heckert, A., & Gondolf, E. (2000b). Predictors of underreporting of male violence by batterer program participants and their partners. *Journal of Interpersonal Violence*, 15, 423-443.
- Heckert, A., & Gondolf, E. (2000c). The effect of perceptions of sanctions on batterer program outcome. *Journal of Interpersonal Violence*, 15, 444-458.

- Quinsey, V., Harris, G., Rice, M., & Lalumiere, M. (1993). Assessing treatment efficacy in outcome studies of sex offenders. *Journal of Interpersonal Violence, 8*, 512-523.
- Raab, S. (2000). Men explode: Have you ever hit a woman? *Journal of Interpersonal Violence, 14* (3), 224-230.
- Rice, M. (1997). Violent offender research and implications for the criminal justice system. *Journal of Interpersonal Violence, 12*, 414-423.
- Roehl, J., & Guertin, K. (2000). Intimate partner violence: The current use of risk assessments in sentencing offenders. *Journal of Interpersonal Violence, 15*, 171-197.
- Rondeau, G., Brodeur, N., Brochu, S., & Lemire, G. (2001). Dropout and completion of treatment among spouse abusers. *Journal of Interpersonal Violence, 16*, 127-143.
- Rosenfeld, B. (1992). Court-ordered treatment of spouse abuse. *Criminal Justice and Behavior, 19*, 205-226.
- Russell, M. (1995). *Court-ordered treatment of spouse abusers*. Thousand Oaks, CA: Sage Publications.
- Saunders, D. (1996). Feminist-cognitive-behavioral and process-psychodynamic treatments for men who batter: Interaction of abuser traits and treatment models. *Journal of Interpersonal Violence, 11*, 393-414.
- Saunders, D., & Hamill, R. (in press). *Journal of Interpersonal Violence*. Washington, DC: National Institute of Justice, U.S. Department of Justice.
- Schare, M. (1992). Needed directions in the psychotherapeutic treatment of drunk drivers. (Paper presented at the 1992 Annual Meeting of the American Psychological Association, Washington, DC).
- Selzer, M. (1971). The Michigan Alcoholism Screening Test (MAST): The quest for a new diagnostic instrument. *Journal of Abnormal Psychology, 76*, 176-181.
- Skeem, J., & Mulvey, E. (2001). Psychopathy and community violence among civil psychiatric patients: Results from the MacArthur Violence Risk Assessment Study. *Journal of Interpersonal Violence, 16*, 358-374.
- Smith, P., & Smith, J. (1999). Beyond the measurement trap. *Journal of Interpersonal Violence, 14*, 177-194.
- Steadman, H. (1982). A situational approach to violence. *Journal of Interpersonal Violence, 7*, 171-186.