

**The author(s) shown below used Federal funds provided by the U.S. Department of Justice and prepared the following final report:**

**Document Title: Supplemental Mental Health Treatment For Batterer Program Participants**

**Author: Edward W. Gondolf**

**Document No.: 223030**

**Date Received: June 2008**

**Award Number: 2003-MU-MU-0002**

**This report has not been published by the U.S. Department of Justice. To provide better customer service, NCJRS has made this Federally-funded grant final report available electronically in addition to traditional paper copies.**

**Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.**

# **SUPPLEMENTAL MENTAL HEALTH TREATMENT FOR BATTERER PROGRAM PARTICIPANTS**

**Edward W. Gondolf, Principal Investigator  
Mid-Atlantic Addiction Training Institute  
Indiana University of Pennsylvania  
Indiana, PA 15705**

**Final Report for  
NIJ Grant #2003-MU-MU-0002 9/01/03-8/31/07**

**Submitted to the National Institute of Justice  
September 20, 2007**

## Abstract

An association between re-assault and mental health problems among batterer program participants has been increasingly documented by research and implied in batterer program guidelines. The majority of guidelines recommend assessment for such problems and referral to mental health treatment in addition to batterer program counseling. However, there is little documentation of assessment results and treatment outcomes. A research project was conducted to investigate the extent of mental health screening, referral compliance, and treatment effectiveness. The project included a formative evaluation of referral implementation, a service-delivery evaluation of the screening and referral, and an outcome evaluation of supplemental mental health treatment on batterer program completion and re-assault of the men's female partners. The *formative evaluation* exposed a few unexpected disruptive events and inconsistencies in referral procedures related in part to administrative turnover and differing priorities among agencies. The implementation modifications resulted in three stages of referral: voluntary referral, transitional referral, and mandatory referral.

The *service delivery evaluation* showed nearly half of the batterer program participants (N=479 of 1043) screened positive on the Brief Symptom Inventory (BSI) and were referred to a local mental health clinic. A concurrent validity test revealed a correlation of the BSI with a more comprehensive screening instrument and a briefer one (n=93), but the test-retest reliability of the BSI was low with a decrease in positive screens over time (n=98). There was also little association with the BSI subscale results and the clinical evaluations obtained by referred men (n=38), and nearly 40% of the clinical diagnoses were for an adjustment disorder not warranting further treatment. Only 30% of the referred men received an evaluation, and 20% obtained some treatment under mandatory referral. Interestingly, at least a third of the referred men acknowledged a need for treatment and was more likely to obtain treatment.

The *outcome evaluation*, based on a 12-month follow-up with female partners (65% response rate; n=308), produced no apparent effect of "intention to treat," represented by mandatory referral, on program completion or re-assault and other abuse indicators. (Re-arrests for violent and other types of crimes were substantially lower for the mandatory referrals in a confirmatory subsample of 300 subjects.) However, there was some preliminary evidence of a "dose response" for evaluated and treated men. Overall, the referral compliance was relatively low, but did increase under mandatory referral reinforced by a system-coordinator and case-manager. Sanctions for non-compliance remained inconsistent and may have affected referral compliance. The results reinforce recent studies exposing the challenges in establishing coordinated community response, and they question the utility of elaborate referral procedures for mental health referral. Alternatives might consider a more simplified referral or integrated services for self-identified problems.

## **Acknowledgments**

This research project was supported by a grant from the National Institute of Justice (NIJ) of the U.S. Department of Justice in Washington, D.C. (NIJ Grant #2003-MU-MU-0002; 9/01/03-8/31/07). The opinions, findings, conclusions, and recommendations expressed in this report on the project do not necessarily reflect the official views of the Department of Justice. Bernie Auchter, Senior Social Science Analyst, NIJ Violence Against Women and Family Violence Program, offered invaluable assistance and guidance in the administration of the project.

The author wishes to thank Candice Petrovich, Chief Executive Officer, and Mark Pudlowski, Chief Operating Officer of the Domestic Abuse Counseling Center (DACC), Pittsburgh, PA, for their assistance in developing and implementing the research project. Also, the DACC office manager, Jennifer Peterson, helped to coordinate the mental health screening at the batterer program and retrieved information on the program attendance and referral compliance. Several DACC counselors conducted the program intake and orientation that included the screening and referral of our research project: Marilyn Arter, Jill Allen Bradley, Joe Carse, and Mark Pudlowski. Jack Simmons, Chief Magistrate of the Pittsburgh Municipal Courts, guided the project's implementation in the courts and provided procedural advice during the course of the research. Marlene O'Leary, the intake coordinator and administrator for the outpatient clinic of Western Psychiatric Institute and Clinic (WPIC), played a key role in developing and administering the procedures for the mental health evaluation and treatment of the referred batterer program participants. Tad Santos offered similar assistance at the mental health clinic associated with Mon-Yough Community Services.

The research project was conducted through the Mid-Atlantic Addiction Training Institute (MAATI) based at Indiana University of Pennsylvania. Special appreciation goes to Crystal Deemer, MAATI Project Director and Administrative Assistant, for her supervision of the screening implementation, data collection, and follow-up interviews. Research assistants, Gayle Moyer, Vera Bonnet, and Tina Gray tracked and interviewed both the men and women in the study. Vera Bonnet also assumed the position of case-manager and system-coordinator during the mandatory referral stage. Nishant Bhattarai assisted ably with data management and analysis, and Megan Kensey worked on data entry and screening scores.

An expert group of researchers offered advice, counsel, and critique throughout the course of the research project and at periodic advisory committee meetings: Edward Mulvey, Director of the Law and Psychiatry Research Program, Western Psychiatric Institute and Clinic, University of Pittsburgh Medical Center; Chuck Lidz, Director of the Center for Mental Health Services Research, Department of Psychiatry, University of Massachusetts Medical School; and Jennifer Skeem, Professor of Psychology and Social Behavior, University of California, Irvine. Several practitioners also served on the advisory committee including Mark Pudlowski, DACC; Marlene O'Leary, WPIC; Lynn Hawker, Women's Center and Shelter of Pittsburgh; Heather Kelly, Assistant District Attorney; and Vera Bonnet and Crystal Deemer, MAATI.

The contributions of these individuals and many others behind the scenes, made this research project very much the result of an extensive team effort and one that represents a wide range of experience and knowledge.

## TABLE OF CONTENTS

### **EXECUTIVE SUMMARY—ii**

### **PART I: INTRODUCTION—1**

**Conflicting Viewpoints—2**

**Previous Research—5**

**Treatment Referral—10**

**Research Questions—11**

### **PART II: RESEARCH METHODS—13**

**Research Design—13**

**Setting—15**

**Sampling Procedures—19**

**Outcome Evaluation—22**

### **PART III: FORMATIVE EVALUATION—31**

**Introduction and Method—31**

**Design Modification—33**

**Implementation Problems—36**

**Underlying Issues—38**

**Recommended Remedies—41**

**Summary—45**

### **PART IV: SCREENING AND REFERRAL—47**

**Introduction—47**

**Sample Characteristics—47**

**Validity and Reliability of BSI Screening—54**

**Batterer Response to Mental Health Screening—67**

### **PART V: BATTERER PROGRAM OUTCOMES—77**

**Introduction—77**

**Results for Program Completion—79**

**Re-assault and Abuse Outcomes—81**

**Confirmation Using Re-arrests—86**

**Summary—88**

### **PART VI: SUMMARY AND DISCUSSION—91**

**Introduction—91**

**Implementation—91**

**Screening—94**

**Referral—97**

**Outcomes—99**

**Limitations and Future Research—102**

### **REFERENCES—105**

### **FIGURES AND TABLES—115**

# **SUPPLEMENTAL MENTAL HEALTH TREATMENT FOR BATTERER PROGRAM PARTICIPANTS**

## **EXECUTIVE SUMMARY**

### **INTRODUCTION**

The relationship between mental health problems and committing domestic violence is a controversial one. Some practitioners in the domestic violence field argue that domestic violence is, at worst, compounded by mental problems but not directly linked to them. Other clinicians assert that domestic violence is more directly linked to psychopathology, and mental problems need to be directly addressed to reduce the reoccurrence of violence. Research on participants in so-called “batterer programs” for domestic violence perpetrators shows men with evidence of psychopathology are more likely to re-assault their female partners, but no distinct profile of psychopathology has been identified. It may be that the psychopathology is directly linked to the violence or that the psychopathology makes perpetrators less responsive to batterer programs.

In either case, there is a growing acceptance that some batterer program participants may warrant supplemental treatment or assistance for mental health problems. The majority of state guidelines for batterer programs currently recommend that batterer program participants be assessed for mental health problems and referred to mental health treatment in addition to attending a batterer program. The guidelines, however, offer little detail on how such an assessment should be conducted, and the practice among batterer programs varies from no systematic assessment to individual clinical evaluations. The prevalence and nature of mental health problems among batterer program participants is, consequently, unclear. Different approaches to assessment using different criteria have produced a range of results. Moreover, there is no empirical evidence that referral for

supplemental mental health treatment is effective in improving batterer program outcomes, and ultimately reducing re-assault.

We conducted a research project to examine the two key practical questions regarding mental health problems among batterer program participants. One, what is the result of assessing for mental health problems and the referral that might accompany it? Specifically, what is the nature and extent of the mental health problems identified in assessment, and what is the compliance and treatment obtained in response to the referral? The second key question has to do with the effectiveness of supplemental mental health treatment in improving batterer program outcomes. Of concern is the effect of referring men to mental health clinics for evaluation and treatment—that is, the “intention to treat,”—as well as the effect of men actually obtaining supplemental mental health treatment—the “dose response.” We would expect intention-to-treat and dose-response to improve the outcomes of batterer program completion and the re-assault of one’s female partner.

In response to several problems in conducting our research project, we also conducted a formative evaluation describing the implementation of assessment and referral. The evaluation examined, in particular, the impact of three unexpected events on the research design and several organization and structural issues among the collaborating agencies. In sum, our research project consisted of a formative evaluation of referral implementation, a “service delivery” evaluation of the results of assessment and referral, and an outcome evaluation of supplemental mental health treatment.

## **REFERRAL IMPLEMENTATION**

### **Screening and Referral Procedures**

The research project was conducted in Pittsburgh, Pennsylvania, and investigated a set of assessment and referral procedures developed among the principle batterer program, domestic violence court, and mental health clinic in the area. The procedures were derived from a series of meetings among representatives from the collaborating agencies and from the lessons of a previous referral project that sent men to a variety of additional social services for educational, parenting, employment, and health needs. To assess batterer program participants for mental health problems, we chose the widely-used Brief Symptom Inventory (BSI). The BSI is a self-administered screening instrument focusing on recent symptoms of distress that are associated with active psychopathology. Its 53-items make the BSI a relatively short instrument to complete and well-suited for the time constraints of program intake. This screening approach to mental health assessment was selected also because of its efficiency and cost-effectiveness. Most batterer programs do not have the resources or staff to conduct individual clinical evaluation, and most rely on a group intake or orientation sessions to bring men into a program. The systematic screening was, moreover, a substantial enhancement at the research site, since mental health referrals were previously made haphazardly by the court based on evidence or testimony at a hearing, or by the batterer program based on staff observations in group sessions.

The screening and referral procedures started with the BSI being administered at program intake, along with the Alcohol Dependence Scale (ADS) and a background questionnaire. The ADS was used because of the high association between alcohol abuse and domestic violence, and because the ADS approximates what is considered a diagnosable mental health disorder neglected in the BSI, namely alcohol dependence and the possibility of a “dual disorder.” At the following orientation



session, the program staff notified the men who screened positive on the BSI or ADS, according to the instrument guidelines, and referred them to one of two mental health clinics for an evaluation and possible treatment. The clinic faxed verification of the referral compliance to the batterer program, and a court liaison from the batterer program reported the compliance with the referral and the batterer program to the court at a court review of the case. The judges were to sanction the batterer program participant further if he was not compliant. We considered these procedures a “mandatory referral” to a mental health clinic in that the batterer program participant was *required* to comply with the mental health referral or receive further sanctions.

The research project was, however, implemented in three stages that established comparison groups for the outcome evaluation. The first stage was “voluntary referral” in which batterer program participants were screened and referred to a mental health clinic but without any requirement to comply or sanctions for non-compliance. This stage represented an “untreated” or quasi-control group for the “mandatory referral” group of batterer program participants. Few men complied with the referral or obtained treatment. In the second stage of “transitional referral,” batterer program participants were required to comply with the mental health referral under threat of further sanctions, but inconsistencies in the mandatory procedures were exposed and modifications introduced. This stage illustrates the screening and referral results without systematic coordination and case-management that tend to be lacking in batterer intervention. Under the “mandatory referral” to a mental health clinic, the inconsistencies and breakdowns were largely eliminated through some modifications in procedures and the introduction of a system-coordinator and case-manager. This stage represented the “intention to treat” condition for the research project.

## **Batterer Program and Mental Health Treatment**

The collaborating batterer program is distinguished by its conventional counseling approach and close linkage to a domestic violence court. The program followed a gendered-based cognitive-behavioral curriculum that conforms to most state batterer program guidelines and the prevailing program models in the field. The domestic violence court conducts a preliminary hearing within a week of arrest, and sends male perpetrators to the batterer program for a minimum of 4 months of weekly group counseling sessions. After program intake and orientation sessions, the men are assigned to one of 20 on-going groups convening throughout the city. The program participation is required as a stipulation of bond and is reviewed periodically by the court. A failure to comply with the program requirements results in jailing, fines, or additional program sessions. Approximately, 10% of the batterer program participants are sent as a condition of probation after prosecution in the criminal courts for more severe aggravated assault or repeated violence offenses.

The mental health treatment was available at one of two mental health clinics affiliated with a major teaching and research hospital. Referrals from the batterer program received the established standard of care for adult outpatients. Upon contacting the clinic, the referred men were evaluated by a clinician, received a diagnosis, and were prescribed appropriate treatment. The treatment generally included up to 12 weeks of individual psychotherapy and the possibility of prescribed medication. Men receiving a dual diagnosis for alcohol dependence and a co-occurring disorder were generally treated in a specialized “dual-diagnosis” unit. An administrator from each clinic was involved in the development of the research project, received the contact calls from the referred men, and assisted with verification of their compliance.

The screening, referral, and treatment together represented an effort towards a “coordinated community response” in which community agencies, including the criminal justice system,

collaborate in a way that reinforces and extends the initial intervention. Our research project, therefore, also provides a test of coordinated community response in terms of implementation and effectiveness.

## **RESEARCH DESIGN**

The research project intended to evaluate the implementation and outcome of screening, referral, and treatment through the three mentioned components: a formative evaluation of implementation, a “service delivery” evaluation of screening and referral, and an outcome evaluation of mental health treatment. Each of these evaluations was conducted through a series of complementary studies with separate methodologies. The outcome evaluation was based on a quasi-experimental design comparing a subsample of men under a mandatory referral to those under a voluntary referral, and also men who actually obtained mental health treatment to those who were referred but untreated. The former comparison represents an “intention to treat” condition and the latter a “dose response” of treatment.

### **Formative Evaluation**

The formative evaluation was based on direct observation of the screening and referral procedures, and training and supervisory meetings. The principal investigator and research assistant also conducted debriefing interviews with administrators, informal conversations with program staff, and discussions with an advisory committee of researchers and practitioners. The investigator and assistant compiled fieldnotes on these activities and periodically compared their summaries to develop conclusions. The main mode of analysis was to compare the conclusions regarding referral implementation to the procedures and protocols established by representatives from the collaborating

agencies. In turn, we attempted to identify the inconsistencies and shortfalls of the implementation, as well as modifications to improve the referral over the course of the research project.

### **Service Delivery Evaluation**

The service delivery evaluation was based on three related studies. First, the results of the screening procedures were assessed. The number and portion of those men screening positive was calculated from 1043 men entering the batterer program during 2004-2006. Also, the characteristics of the “positive” men who were subsequently referred (N=479) were compared to the men who screened negative and were not referred for supplemental mental health treatment. This comparison was to help describe and distinguish the referral sample.

Two, reliability and validity tests were conducted with the BSI screening instrument. These tests were in part a response to the unusually high portion of positive screens identified in the previous study. “Test-retest reliability” was examined by comparing the scorings on the initial BSI screening conducted at program intake to the BSI scorings administered at one month to five months later at either a program counseling session or follow-up interview with the men (n=98). “Concurrent validity” was examined by comparing scoring results of the BSI with two other screening instruments used to identify mental health problems and psychiatric disorders: the more comprehensive Psychiatric Diagnostic Screening Questionnaire (PDSQ) and the shorter Personality Assessment Screener (PAS). The instruments were administered to a subsample of 93 men following batterer counseling sessions at 7-10 weeks after program intake. We considered the “predictive validity” of the BSI as well. Cross-tabulations, logistic regressions, and receiver operating characteristics (ROC curves) were computed to determine the association of the nine BSI subscales and global index with batterer program completion and partner re-assault. (Re-assault was based on interviews with the

men's female partners during a 12-month follow-up period, as explained further under "Outcome Evaluation" below.)

Three, compliance with the mental health screening was examined using three sources of data: case-management interviews conducted 3-4 weeks after program intake with the men referred under mandatory referral (n=33 interviewed), the 5-month follow-up debriefing interviews with the referred men (n=254 interviewed), and clinical records for all the men referred under the mandatory referral (n=148). The extent of compliance to scheduling an evaluation appointment, obtaining an evaluation, and receiving treatment was tabulated, along with the clinical diagnosis and treatment length. Also, the men's self-identified needs and symptoms, their expectations for treatment, and the benefits from treatment were summarized from the men's interview questions. Drawing on characteristics from the background questionnaires and BSI screening at program intake, we also attempted to identify predictors for referral compliance using logistic regressions.

## **Outcome Evaluation**

The outcome evaluation was the most extensive and complex component of our research project. To determine the effectiveness of the "intention to treat," the batterer program outcomes were compared for the three referral stages (N=479), and to examine the effectiveness of a "dose response" to treatment—that is, actually receiving treatment—the outcomes were compared for a) the men who obtained an evaluation or received treatment and b) the men who did not comply to the referral under the mandatory referral stage (n=148). The comparisons were computed using cross-tabulations of the treatment condition with the outcomes, and confirmatory logistic regressions controlling for background characteristics (even though the characteristics across the comparison

samples were equivalent). Re-arrests rates for a stratified random sample (n=300) were also analyzed as a further confirmation of the re-assault results.

The principal outcome measures were batterer program completion and re-assault during a 12-month follow-up period. “Batterer program completion” was derived from the computerized attendance records maintained by the batterer program with “completion” identified as satisfactorily attending the minimum requirement of 16 weekly sessions. “Re-assault” was measured as any physically aggressive tactic committed against one’s female partner during the follow-up period. Several additional indicators of abuse were also used, including threats, severe re-assault, and the women’s perception of overall well-being and safety. Research assistants interviewed the men’s initial and new female partners by phone at program intake and every three months over the 12-month follow-up period. The response rate for the full 12-month follow-up was 65% (n=308 of 479).

## **RESULTS**

### **Referral Implementation**

Our formative evaluation identified three unexpected events that interrupted the initial experimental design and forced a modification to the quasi-experimental design of comparative subsamples. The district attorney objected to the random assignment of an experimental design because of a high-profile murder by a former batterer program participant. An administrative scandal at the batterer program led to the dismissal of the program director and assistant director under accusations of embezzlement and fraud. The “institutional review board” of the collaborating clinics questioned the ethics of withholding mental health treatment from some men who appeared to warrant it. These events exposed some of the “real world” circumstances that not only make experimental evaluation difficult, but also can interfere with community coordinated response.

Batterer programs and program context remain vulnerable to administrative disruptions that are hard to anticipate.

We also identified inconsistencies and breakdowns associated with nearly every step of the screening and referral process. For example, some men were not properly notified about referral instructions because of staff or participant absences, and compliance verification from the clinics to the batterer program was sometimes not completed. Procedural modifications corrected these and other problems, but the court sanctions for non-compliance remained inconsistent even during the mandated referral stage. Several underlying issues contributing to the initial inconsistencies were also exposed. Of particular note were the administrative absenteeism and turnover, administrative-staff gaps, client overload, and differing agency priorities. These issues reinforce the challenges facing coordinated community response that have surfaced in other recent studies of batterer intervention implementation.

## **Service Delivery**

The service delivery evaluation showed a substantially high portion of batterer program participants warranting a mental health referral according to screening criteria. However, a relatively low portion and perhaps exceptional group of these men received mental health treatment, even under mandatory referral. More specifically, over 40% screened positive on the BSI; scorings on the ADS and direct referral from judges increased the referrals to nearly 50%. The “positive” screens appeared more at risk for program dropout and re-assault than the negative screens, according to a comparison of background characteristics, and more in need of an expanded intervention. Only 30% of these men obtained an evaluation, and 20% of the referred men eventually received treatment under the mandatory referral. A portion of these “treated” men were already in treatment prior to the

referral from the batterer program, and less than 5% of the referred men attended 10 or more treatment sessions.

These results translate into less than 10% of the total number of batterer program participants receiving supplemental mental health treatment. However, the portion of compliant men dramatically increased over voluntary referral with less than 7% complying (including men already in treatment). Interviews with the men also revealed an increase in compliance, especially following contact from a case-manager and after a few months from program intake. Enhanced implementation toward increased consistency, accountability and supervisions did appear to improve compliance, as it has with batterer program attendance.

The reliability and validity tests suggest that the small portion of treated men is not necessarily attributable to over-screening with the BSI. On the one hand, the BSI scorings were highly correlated to those with the PDSQ and PAS. On the other hand, those scoring positive on the BSI retest decreased substantially, but this may be related to a change in men's circumstances or lessons from the batterer program, according to interviews with the men. Moreover, the predictive validity of the BSI subscales was weak. The most disconcerting finding was that nearly 40% of the referred men who did receive a clinical diagnosis (n=38) were identified as having an adjustment disorder and did not warrant further treatment. The BSI subscales, such as depression, did not, as well, match with the clinical diagnoses.

Interestingly, the men reported a sharp decline in compliance across the referral steps from scheduling an appointment to attending treatment sessions. At the same time, at least a third of the batterer program participants initially indicated that they would benefit from mental health treatment, and identified symptoms that might warrant such treatment. These expectations did approach significance in a regression predicting compliance along with probation supervision and a protection



order. The BSI scorings were not predictive of receiving treatment. The vast majority (86%) who did receive treatment reported that it was helpful to some or a great extent not only in reducing their symptoms but also in improving their batterer program attendance. Despite the low compliance overall, 88% of the men agreed that the mental health referrals should be continued.

## **Treatment Outcome**

We found some tentative evidence that supplemental mental health treatment might improve batterer program outcomes. As for intention-to-treat, the completion rate for the referred men was only slightly lower than that for the men not warranting referral (57% vs. 62%), and the re-arrest rates were similar for the two groups (any crime: 23% vs. 29%) despite the expectation that the referred men would have poorer outcomes because of their additional problems. (Re-assault information was not obtained for the non-referred men.) Also, the re-assault rate of 32% during the 12-month follow-up did not significantly differ across the referral stages (e.g., voluntary and mandatory referral), nor did indicators for other forms of abuse or the women's well-being. The results for program completion and re-assault were confirmed in the logistic regressions controlling for background characteristics. The women's sense of safety and likelihood of being hit were, however, significantly lower during the mandatory referral stage, while re-arrests for violent crimes and crimes in general were 40% less likely during the mandatory stage compared to the voluntary.

Regarding dose-response, the men obtaining a clinical evaluation showed a statistically significant increase in batterer program completion (76% vs. 38%;  $p < .05$ ), as did the treated men (68% vs. 45%;  $p < .05$ ). The evaluated men also were a third less likely to re-assault their partners (22% vs. 31%; n.s.) as were the treated men (19% vs. 30%; n.s.), but these differences were not statistically significant with the small numbers of treated men. Other abuse indicators tended in the

same direction for the evaluated men but were similar for the treated men. The re-assault tendency was confirmed when controlling for batterer program completion, and both the program completion and re-assault results were consistent when controlling for background characteristics.

In sum, while the intention-to-treat represented by the mandatory referral did not significantly improve outcome overall, the dose-response showed a substantial tendency in the expected direction. That is, men who obtained a mental health evaluation and men who received treatment were more likely to complete the batterer program and less likely to re-assault their female partners. This latter tendency must be viewed with extreme caution because it is not statistically significant and based on a very small number of compliant men under mandatory referral. It is, furthermore, likely to be confounded by motivational differences and other characteristics that were not controlled in the comparison between those men treated and not treated. Even if the apparent treatment effect were replicated in a controlled comparison, the effect size would be very small overall.

## **DISCUSSION**

Screening and compliance for supplemental mental health treatment produced relatively low compliance even under a mandatory referral system. Some preliminary evidence did suggest that treatment was beneficial for the small portion of perhaps exceptional men who did receive it. The mandatory referral may not, however, be worth the additional layer of collaboration, cost, and staff to implement mandatory referral given the relatively small and still uncertain effect. This may especially be the case given that the relationship between mental health problems and batterer program outcomes remains in question, or at least complex.

At the same time, the substantial increase in compliance under mandated as opposed to voluntary referral is worthy of note. It suggests that coordinated community response helps to

increase service delivery, but a system-coordinator and case-manager seem essential to establishing consistency in procedures. Our research project, however, confirms the challenges in implementing coordinated community response and the incomplete implementation even with extra resources and staff.

The referral compliance may have been increased if the mandatory referral were more completely implemented, namely with more consistency in the court response to non-compliance. The referral may also be more efficient and treatment more clearly effective under a different screening procedure--one that is more selective and considers the men's self-identified needs and motivation. We found, for instance, the men who saw some benefit in mental health treatment to be more likely to seek and receive it. Moreover, a simplified referral system would help reduce inconsistency in implementation and the progressive non-compliance over referral steps. Programs that integrate batterer counseling and mental health treatment are one possibility in this regard, but do raise the likelihood of additional costs and the diffusion of domestic violence education.

Interestingly, the BSI subscale for hostility was the most highly associated with negative outcomes, and resistant and uncooperative men were more likely to receive a "default" clinical diagnosis of adjustment disorder and not be recommended for treatment. This may in part represent a clash between criminal justice and mental health priorities and approaches. That is, the criminal justice system relies heavily on coercion and sanctions and the mental health clinics on voluntary and motivated help-seeking. The men with the most severe mental problems were, moreover, not necessarily the ones to receive treatment. In sum, the men most in need of expanded intervention may have avoided it. More obviously needs to be done to identify and contain those batterer program participants most at risk for re-assault and harming their partners.

## **PART I: INTRODUCTION**

A mounting movement within the domestic violence field is bringing more attention to the psychopathology of domestic violence offenders, commonly referred to as “batterers.” With it has come controversy and even conflict over the role of mental health treatment for batterers between clinical psychologists and many battered women’s advocates and batterer counselors working with them. The lead chapter of the book *Current Controversies on Family Violence* (O’Leary, 1993) outlines this controversy and concludes that psychopathology is often a compounding or reinforcing factor in a batterer’s violence that warrants specific treatment. Moreover, a recent journal article accentuates the controversy with its severe criticism of conventional batterer programs’ failure to adequately consider the psychological problems associated with violence (Dutton & Corvo, 2006). Two recent conferences, convened by battered women’s advocates, illustrate the cautions in this regard. The Violence Against Women Office of the Illinois Coalition Against Domestic Violence held a critical conference in November 2002 entitled “The medicalization of domestic violence.” National speakers from the domestic violence field raised questions about the utility of mental health treatment for batterers, as well as the consequence of mental health treatment for battered women. In October 2002, the Georgia state coalition sponsored a conference for its batterer program staff and counselors. A contentious debate emerged in the discussion and summary sessions over the role and effectiveness of mental health treatment for batterers.

There is at least some preliminary research suggesting that mental health disorders are associated with re-assault and more severe abuse by batterer program participants, as the *Controversies* chapter argues. Many state standards, moreover, require batterer programs to assess the mental health of program participants, but the assessment or treatment is vague for the most part.

These standards typically indicate, however, that any mental health treatment must be in conjunction with batterer treatment and not a replacement for it. Moreover, batterer programs are increasingly developing their own assessment protocols and integrating treatment for mental health problems. There remains, however, a lack of empirical evaluations of the contribution of mental health treatment on batterer program outcomes. In sum, we need some indication of the mental health treatment that batterers might warrant, and of the effectiveness of such treatment in improving batterer program outcomes.

## **CONFLICTING VIEWPOINTS**

### **CLINICAL PSYCHOLOGISTS**

From the point of view of many clinical psychologists, domestic violence is related to psychopathology. If not an extension of psychopathology, it is at least substantially reinforced by it (see O'Leary, 1993; Dutton & Corvo, 2006). Individual psychopathology may, in fact, be the main reason that some men, living in similar social environments, are assaultive and others are not. The assumption appears to be that violence is compounded or reinforced by some underlying psychiatric disorder or problem. The intrapsychic pain of depression, for instance, may be expressed and perhaps eased in a violent outburst, or the misperceptions of paranoia or delusions might contribute to conflict and aggression. Such disorders could be a reaction to childhood trauma, attachment issues, or borderline tendencies that have been ascribed to many batterer program participants (e.g., Dutton, 1998; Stosny, 1995).

In this view, violence avoidance techniques taught in conventional cognitive-behavioral counseling are insufficient, and the confrontation of social sanctions and personal responsibility could exacerbate underlying emotional pain and frustration. Some clinical psychologists working in the

domestic violence field have recommended, therefore, more of a psychodynamic approach to batterer counseling in order to address the underlying mental health issues (e.g., Dutton, 1998; Stosny, 1995). More generally, there is a call to systematically assess the mental health needs of batterers and refer them accordingly to supplemental treatments that may include individual psychotherapy or addiction treatment (e.g., O’Leary, 1993; Stordeur & Stille, 1989). Having, at minimum, the batterer counseling focused on behavior seems a basic safety precaution, and even the more psycho-dynamic approaches in batterer programs include those. The overall assumption is that addressing the psychopathology of batterers will reduce the men’s violence and lower other forms of abuse as well.

## **BATTERED WOMEN’S ADVOCATES**

Battered women's advocates and victim rights proponents tend to view domestic violence as rooted primarily in learned or socialized behavior and reinforced by cultural norms (Yllo, 1993). Domestic violence is, more specifically, an extension of the “power and control” that men tend to exert over women in our society. Advocates have consequently pressed for decisive criminal justice sanctions to domestic violence offenders that would “teach” them that violence is not acceptable, as well as help contain them from continued violence. This sort of response has been expressed in mandatory arrests for domestic violence and full-prosecution to hold batterers accountable for their behavior. It has resulted as well in a gender-based cognitive-behavioral model of group counseling for batterers that attempts basically to help men “learn” alternatives to violence (see Pence & Paymar, 1993).

Battered women's advocates tend, therefore, to view psychopathology not only as secondary to violence, but also as a possible distraction from the real issues (Bancroft, 2002). They believe that the focus needs to be on the social acceptance and reinforcements of violence and on men’s

rationalizations and justifications for domestic violence (e.g., Russell, 1995). Psychopathology, from this point of view, inadvertently diverts us from the social action that holds individual men accountable. It may even reinforce the years of excusing domestic violence as peculiar to an exceptional subgroup of deviants. Advocates also draw on experience and documentation that mental health services have tended to neglect violence and safety issues in the course of treatment, as studies of mental health treatment have shown (see Gondolf, 1998). Moreover, psychopathology would warrant a more central role for clinical psychologists in the assessment and treatment of domestic violence, and possibly displace the experience and input of battered women's advocates.

## **INTEGRATIVE APPROACH**

A more integrative approach to batterer intervention and batterer programs is also emerging within the field (e.g., Hamberger, 2002; Pettit & Smith, 2002). Researchers and practitioners alike are accepting a more multi-faceted view of domestic violence and the need for a more comprehensive intervention, as a recent review article on the trends in batterer programs promotes (Mederos, 2002). What might be termed "conventional" batterer programs are increasingly identifying and addressing mental health, substance abuse, and childhood trauma issues (Gondolf, in press-b). A national conference on batterer intervention, "From Roots to Wings: The Future of Batterer Intervention," openly confirmed the need to broaden the cognitive-behavioral focus of batterer programs and asserted it as the current direction (Dearborn, Michigan, in November 2006; see [www.biscmi.org/documents/biscmi10thconference.html](http://www.biscmi.org/documents/biscmi10thconference.html)).

There is admittedly a diversity of ways and positions on how best to accomplish this more integrative approach, and the degree of attention and resources it warrants. Some batterer programs require an extensive individual evaluation for each man that enters batterer group counseling; others

rely on a group screening and observation to identify men that may warrant additional services and supervision. In some domestic violence courts, psychological and substance abuse evaluations are ordered by the court based on the circumstances of the case, and men must complete the evaluations in addition to the batterer program. Men identified with additional problems may be referred to mental health clinics for supplemental treatment, treated in separate sessions by batterer program staff, or given special attention within the batterer group counseling.

Conflicts in orientation and emphasis still beset the field, but these may represent more the extremes rather than the emerging convergence (see for example, Dutton & Corvo, 2006). The experience and observations of this researcher are that the differences are historically rooted. Conventional batterer programs were developed largely in response to a societal critique and safety concerns of battered women's advocates often in contradistinction to clinical psychologists and family therapists who were viewed as neglecting violence against women. The evolution toward a coordinated community response has brought cross-training and service collaborations that have helped to bridge a variety of services and perspectives (Pence & Shepard, 1999). There is little research or evaluation, however, that demonstrates the implementation and effectiveness of the integration, especially in terms of addressing the mental health problems of domestic violence perpetrators.

## **PREVIOUS RESEARCH**

### **MENTAL DISORDERS AND VIOLENCE**

A long line of research with community and general population samples has shown a higher incidence of violence among those with mental health disorders. The inference of some researchers and clinicians is, therefore, that we can expect a high prevalence of mental health disorders among



clinical samples of domestic violence perpetrators (e.g., Dutton & Corvo, 2006). As early as 1986, Bland and Orn (1986) assessed a random sample of an adult urban population (n=1200) using the Diagnostic and Statistical Interview Schedule and found that those with one of three major diagnostic categories (i.e., antisocial personality, major depression, alcohol dependence) were twice as likely to have abused a family or nonfamily member in the previous year than those adults without a diagnosis. Other researchers have shown that the mentally ill do not pose a high risk in absolute terms. Only about 7% of those with major mental disorders (but without substance abuse) engage in any assaultive behavior in a given year (Swanson, 1994), and only about 20% of those who heavily drink in a given year (Kantor & Straus, 1987).

More recent evidence from longitudinal population studies indicates, however, that family violence is associated with a broad array of psychological and social factors during one's childhood and adolescent development (Ehrensaft, Moffitt, & Caspi, 2004; Moffitt, Caspi, Rutter, & Silva, 2001). Additionally, studies drawing on the National Comorbidity Survey have shown that a fairly stable portion of the population shows evidence of a mental health disorder—about 30% and only about a third of these receive some form of treatment. Recently, psychiatric researchers have identified intermittent explosive disorder (IED) as responsible for the violence of as much as 30% of domestic violence perpetrators (Kessler, Coccaro, & Fava, 2007). Conventional cognitive-behavioral batterer programs may, however, be appropriate and adequate for IED (Gondolf, 2006).

Some fundamental issues make it difficult, however, to extend findings from general population samples to clinical samples (e.g., men in a batterer program). Only about 1-2% of those men who are physically aggressive toward their partners within a year are arrested and sent to a batterer program (based on estimates from Tjaden & Thoennes, 2000). This small percentage is not necessarily representative of the general population, especially given the social factors related to

severe violence and being arrested. Moreover, the relationship of mental disorders to violence is a complex one, as psychiatric studies with violent patients have shown (e.g., Monahan et al., 2001). The violence may be “associated” with a certain disorder or merely a “complication” of it. We are left, therefore, with the practical question of whether the prevailing cognitive behavioral batterer programs adequately or effectively reduce the violence of men with mental health problems, and whether supplemental treatment that specifically addresses the additional problems is warranted.

## **MENTAL DISORDERS AND BATTERERS**

There has been a substantial amount of research investigating the personality and psychological characteristics of batterers. Several of these studies have used the Millon Clinical Multiaxial Inventory (Millon, 1994) to identify psychopathology among batterers (i.e., DSM-IV Axis I disorders or DSM-IV Axis II disorders). This research suggests a substantial prevalence of major disorders and their association with batterer program dropout and re-assault (Hamberger & Hastings, 1991; Hamberger et al., 1996; Gondolf, 1999b). Other research addressing distinguishing personality traits, batterer types, and risk assessment have also indirectly raised questions about psychopathology (e.g., Dutton, 1998; Holtzworth-Munroe & Stuart, 1994). The two most violent batterer types, for instance, are in part differentiated by depressive/compulsive and narcissistic/antisocial tendencies. The prevailing risk assessment instruments (i.e., SARA and DA) include indicators for major depression and thought disorder, or at least suicidal threats and attempts (see Dutton & Kropp, 2000; Roehl & Guertin, 2000).

A few of these studies specifically indicate a high prevalence of psychopathology (i.e., DSM-IV Axis I disorders or DSM-IV Axis II disorders) among batterer program participants (Hamberger et al., 1996; Gondolf, 1999b). According to studies with the MCMI-I, as much as 40% of men entering

programs show evidence of psychopathology, much of which might be expressed in the form of personality tendencies or disorders (Hamberger & Hastings, 1991; Hamberger et al., 1996). The vast majority of the personality disorders, according to the MCMI interpretative manuals (Choca & Van Denburg, 1997), would most appropriately be addressed by highly structured cognitive-behavioral treatment much like those commonly used in batterer counseling (White & Gondolf, 2000). In our previous studies with the MCMI-III, a quarter of the men in four cities (N=854) appeared to have Axis I disorders (i.e., the MCMI “clinical syndromes” or “severe syndromes”), excluding alcohol dependence and anxiety disorders, and as much as 40% had evidence of an Axis I disorder when anxiety disorder and alcohol or drug dependence were included (Gondolf, 1999b). The main limitation with this line of research is that the MCMI generates profiles of a variety of personality dimensions that may approximate DSM-IV disorders but do not directly match them. The MCMI is not a DSM-IV diagnostic tool in itself.

The question remains whether the mental health disorders are related to re-assault and especially to continuous or extremely harmful assaults. Batterer program completion (of a minimum of 3 months) is substantially lower among men appearing to have an Axis I disorder on the MCMI (49% vs. 67%;  $X^2[1]=9.93$ ;  $p>.001$ ). Those with severe pathology are also a third more likely to re-assault their partners (29% vs. 41%;  $X^2[1]=7.96$ ;  $p>.01$ ), and twice as likely to repeatedly re-assault during the 15-month follow-up (8% vs. 17%;  $X^2[1]=9.93$ ;  $p>.01$ ). We used a specialized regression model (i.e., Generalized Estimating Equations) to control for demographics and previous behavior, and to account for program dropout and longitudinal measures. In this analysis, severe psychopathology doubled the risk of re-assault during the follow-up (Jones & Gondolf, 2001). The only other significant predictor, among demographic, behavioral, and other personality measures, was a history of non-domestic violence arrest. In a multinomial logistic regression for multiple outcomes

including repeated re-assault; none of the personality, mental health variables, or various classifications of batterer types were significant predictors of re-assault or repeated re-assault (Heckert & Gondolf, 2005). This finding leaves questions about the impact of mental health disorders on severe re-assault, but does not deny the possibility that the effects of psychopathology might be indirect, as mediated by conditional factors, in some underlying causal model.

The most striking finding in the longitudinal analysis was that drunkenness during the follow-up was the most influential predictor of re-assault, increasing the likelihood of re-assault by four (Jones & Gondolf, 2001). Men who were drunk on almost a daily basis, according to their partner's reports, were 16 times more likely to re-assault. Measures of drinking frequency and alcohol dependence at program intake were not significantly associated with the re-assault outcome (e.g., Michigan Alcohol Screening Test, Selzer, 1971; MAST > 4). Alcohol and drug treatment in the previous follow-up interval reduced the probability of re-assault by roughly 30-40%, although this effect was not statistically significant at conventional levels. The failure of this effect to achieve statistical significance might be due in part to insufficient measurement, since the treatment measure did not account for extent or nature of treatment.

The alcohol finding suggests that severe addiction, which is unresponsive to treatment, should be of great concern to batterer programs. At least 45% of these cases were compounded with co-occurring Axis I diagnoses according to the MCMI. A three-way cross-tabulation (re-assault x drunkenness during follow-up x Axis I disorder on the MCMI) shows that these "dual-diagnosis" cases were four times more likely to commit severe re-assault (according to the Conflict Tactics Scale; Straus, 1979), as opposed to twice as likely if drunk without an Axis I disorder. The trend in the alcohol treatment field has been to separate dual-diagnosis cases for specialized treatment since they tend to be unresponsive to conventional addiction treatment and exhibit problematic behavior.

The multi-site evaluation of patient-matching funded by the National Institute of Alcoholism and Alcohol Abuse (NIAAA; Project MATCH, 1997), for instance, found that the addiction patients with “severe psychiatric disorders” did not perform well in any of the 16 treatments being tested.

## **TREATMENT REFERRAL**

Across the country, state standards for batterer programs have raised the issue of mental health problems but are generally vague on how they should be addressed. Our review of the batterer program standards available on-line ([www.biscmi.org](http://www.biscmi.org)) showed that 89% of the 36 states currently with standards recommend some consideration of mental health problems during program intake. This mention varied from formally assessing psychiatric disorders or mental health status (49% of the states) to merely noting psychiatric or mental health history (35%). However, only 59% of the state standards specifically designate a treatment response to these problems, and that varies from screening out men with severe disorders to referral to mental health clinics. There is a prohibition or caution against mental health treatment in place of batterer counseling throughout the standards.

A few states, such as Florida and Washington, do require an individual psychological assessment for all men convicted of domestic violence and concurrent mental health treatment when appropriate. The prevailing problem here is that referrals are not consistently made, treatments vary widely, and compliance is not monitored. There is no unified summation of the clinical diagnoses and prescribed treatments applied to court-ordered batterers. Moreover, the apparent inconsistency of, and non-compliance to, mental health treatment makes it difficult to assess treatment utility and impact. The contribution, if any, to batterer counseling outcomes and the safety of women, more specifically, remains undocumented and unclear. The controversy over mental health treatment intensifies without some empirical basis to sort out the assumptions and claims. It is especially

important to address mental health disorders if they are related to batterer program outcomes, as the previous, albeit limited, research suggests.

## **RESEARCH QUESTIONS**

The first and most fundamental need is to implement and assess a screening and referral process for supplemental mental health treatment—that is, mental health evaluation and prescribed treatment in addition to attending a batterer counseling or educational program. This step would begin to address the minimum recommendation of state standards and also pose a middle-ground response to the controversy regarding mental health treatment for domestic violence. What portion of batterer program participants would warrant referral under conventional screening protocols? What is the nature of their mental health symptoms, problems, or disorders? What is the men’s response in terms of referral compliance and treatment received? In sum, what are the results of screening and referral in terms of “service delivery”?

A second and more crucial need is to test the effectiveness of supplemental mental health treatment. Does it improve batterer program outcomes in terms of reducing program dropout and re-assaults of one’s female partner? What is the additive effect to conventional batterer programming both in terms of the intention-to-treat and actually receiving treatment (i.e., dose-response)? In other words, does screening and referral improve outcomes (i.e., intention-to-treat), and, if not, is there evidence that doing more to ensure treatment produces an improvement (i.e., dose-response). It would also be helpful to determine which men are the most responsive to such treatment—what screening scorings or clinical diagnoses are associated with receiving treatment and improved outcomes?

Research on these questions might initially address psychopathology, or mental health problems, in terms of DSM-IV Axis I disorders (e.g., major depression, bipolar, delusional, posttraumatic stress, anxiety disorders), including dual-diagnosed addiction. Axis II personality disorders may also be of concern but the most prevalent of these disorders among batterer program participants (i.e., antisocial, narcissistic, and passive-aggressive disorders) are generally prescribed the kind of structured, cognitive-behavioral counseling used in the majority of batterer programs (Rice, 1997; White & Gondolf, 2000). Moreover, some clinical researchers suggest that many of the extremely antisocial men are unresponsive to counseling in general, and typically do not comply to batterer counseling, let alone supplemental mental health treatment (Dutton, 1998; Jacobson & Gottman, 1998). The focus on Axis I disorders, therefore, would treat those most in need of additional treatment and most likely to benefit from it.

The effort to screen, refer and treat batterer program participants for mental health problems represents, moreover, an answer to the call for “coordinated community response” in the domestic violence field (e.g., Pence & Shepard, 1999). Coordinated community response refers to the collaboration and cooperation among social services and agencies in a way that expands and reinforces intervention in domestic violence cases. The sanctions and oversight from particularly the criminal courts—what is often referred to as “accountability”—assume a central component of this sort of response. They help to coerce especially resistant batterers to comply with service referrals and to contain men who are not responsive to the intervention. Research on supplemental mental health treatment might, therefore, be considered a test of coordinated community response as well as of treatment effectiveness by itself.

## **PART II: RESEARCH METHODS**

### **RESEARCH DESIGN**

To answer these basic questions, we conducted a research project on the “service delivery” of screening and referring batterer program participants to mental health treatment, and on the outcomes of receiving such treatment. We first assessed the “service delivery” in terms of the portion of men who screened positive for referral, the characteristics and profiles of the referred men, the extent and nature of their compliance to referral, and the men’s response to referral. Our initial findings also led us to examine the validity and reliability of the screening and possible alternatives to the screening approach. Moreover, we encountered inconsistencies and breakdowns in the referral process that prompted a formative evaluation to better understand the implementation of referral and help qualify and interpret the outcome evaluation.

Secondly, we conducted the outcome evaluation by comparing the batterer program completion and re-assaults against one’s female partner for three subsequently recruited subsamples of referred program participants: 1) a subsample of men voluntarily referred to supplemental health treatment which represented a “no-treatment” or quasi-control group, 2) a subsample of men under a mandatory referral, including oversight from a case-manager which represented the “experimental” or treated group, and 3) a subsample of men referred during a transitional period in which the mandatory referral was not consistently implemented or enforced. As our implementation demonstrated, only a very small portion of men referred on a voluntary basis (i.e., with no court sanctions or consequences if they did not comply) received treatment, or even made an appointment for a clinical evaluation. The transitional subsample, moreover, enabled us to consider the service delivery and outcome over a gradation of referral implementation and requirements. This subsample



also is likely to represent the “real world” conditions of most batterer program referral, where the ideal extent and consistency of a court-supported mandate is not feasible.

We initially planned an experimental evaluation comparing the outcomes of conventional batterer counseling supplemented with mental health treatment to batterer counseling alone. Batterer program participants who screened positive for referral were to be randomly assigned to the two treatment options. If properly implemented, this design would provide a controlled test of the effectiveness of supplemental mental health treatment, or at least the *intention* to treat the men assigned to it. For reasons discussed in “Part III: Formative Evaluation,” this design was replaced by the quasi-experimental design comparing three subsamples sequentially recruited as more stringent referral requirements were established.

Our sample of referred program participants was drawn from the population of court-ordered men to 16 weeks of required batterer counseling in Pittsburgh (N=1043). Over a two year period beginning in 2004, 479 men were positively screened for mental health problems at batterer program intake, and referred to one of two collaborating mental health clinics for an evaluation and appropriate treatment. For the assessment of service delivery, the screening results were tabulated, a debriefing interview was conducted with the men, information was collected during the case-management during the mandatory referral, and clinical records were obtained from the mental health clinics. For the outcome evaluation, the men’s female partners were interviewed by phone at program intake and every three months over a 12-month follow-up in order to determine the extent of re-assault. Batterer program completion was determined from computerized attendance records available from the program.

This research project actually represents a series of studies addressing the overall objectives. Each study is therefore separately presented in two parts summarizing the results: “Part IV: Screening

and Referral” that focuses on the service delivery, and “Part V: Batterer Program Outcome” that examines the impact or effectiveness of the supplemental mental health treatment. These parts include the rationale, methods, and results of each study, rather than this methods section compile all the separate studies together. We believe that the results of the studies are easier to assess when they directly follow the methods that produce them. The results of the separate studies are then summarized, integrated, and qualified in “Part VI: Summary and Discussion.” Qualifications, implications, and recommendations conclude that final part.

In this part of the report, the setting of the research project is described along with the referral system being investigated. This description of setting and system enables a comparison to other intervention programs and systems that might indicate their representativeness, as well as exposes features or components that might help in interpreting the outcomes. The procedures for subject recruitment are also outlined and the methods for the outcome evaluation are discussed in some detail. The outcome evaluation is the most extensive and complex study within the overall research project. It involves a series of follow-up interviews with the female partners of the batterer program participants to assess the outcome in terms of re-assault. Also, the analysis of the relationship of supplemental treatment to batterer program outcome requires some elaboration and is therefore presented below.

## **SETTING**

### **BATTERER COUNSELING PROGRAM**

The setting for the research project was the Domestic Abuse Counseling Center (DACC) in Pittsburgh, Pennsylvania. DACC offers weekly sessions of 1½ hours to groups of 13-15 men for a required duration of 16 weeks. A staff counselor follows a primarily instructional or didactic

approach that conforms to the gender-based cognitive-behavioral curriculum prevalent in the field (e.g., Pence & Paymar, 1993; Stordeur & Stille, 1989), and consistent with most state program guidelines (Austin & Dankworth, 1999). A structured curriculum is presented in a manual available on-line and on CD-rom, and monitored bi-monthly by a clinical director. The main topics include the nature and impact of abuse, the consequences and costs of abuse, taking responsibility for one's abuse, ways to avoid abusive behavior, and beliefs and attitudes that sustain abusive behavior.

Program involvement begins with a group intake session at one of four locations across the city. Program staff collects necessary background information from the men, review the program and its policies, and present abuse and violence avoidance skills. The following week is an orientation session at the same location. The staff further discusses the avoidance skills and program expectations, and then assigns each man to one of 20 counseling groups nearest his home. A weekly fee is set on a sliding scale (\$5-\$50 per session; 37%<\$20; 32%=\$20; 31%>\$20). Two unexcused absences, delinquent payments, or an arrest for re-assault result in the man's dismissal from the program and his return to court.

The participants are ordered to the batterer program at a preliminary hearing in a specialized domestic violence court as part of a bond stipulation. The court procedures ensure a "swift and certain" response to domestic violence arrests. Not only are the men in court within 10 days of arrest, but they are also subject to periodic oversight from the court. The men must demonstrate compliance to the batterer program at a court review hearing 30 days after the initial referral and again at 90 days. The program office staff compiles reports of non-compliant men and presents them to the court via a program court liaison. The court liaison is present at each court session to receive men ordered to the batterer program and present compliance information for participants under review. The court issues bench warrants and additional sanctions (i.e., fines, extended counseling, jailing) to the non-

compliant men. This “judicial oversight” is in contrast to the slow and uncertain response of the vast majority of probation offices to non-compliance (see Gondolf, 1999b). In addition, approximately 10% of the batterer program participants are sent to the program by the city probation office. These are men who were forwarded to a higher criminal court because of severe violence or previous offenses.

The batterer program has been evaluated extensively, showing an effect in reducing the rate of re-assault over court intervention only. As part of a prior multi-site evaluation, “instrumental variable” structural equations and propensity scores were used to identify a moderate “program effect”—program completion reduced the likelihood of re-assault by 50% (Gondolf & Jones, 2001). Also, a previous comparative evaluation showed that the three-month Pittsburgh batterer counseling was as effective as longer six-month and nine-month programs with similar curriculum approaches (Gondolf, 1999b). This equivalent effectiveness was in terms of re-assault, calls to the police, the women’s perception of safety, and the women’s quality of life.<sup>1</sup> (The Pittsburgh program is currently four-months in length.)

## **MENTAL HEALTH OUTPATIENT CLINICS**

The mental health treatment was provided by one of two outpatient clinics affiliated with the University of Pittsburgh Medical Center (UPMC). Western Psychiatric Institute and Clinic (WPIC), a branch of UPMC, administers and coordinates these clinics. WPIC is a large, urban teaching hospital, known internationally as a research and clinical center on the cutting edge of psychiatric care. The two clinics provide general outpatient services, as well as specialized services for patients with co-occurring Axis I and substance use disorders. The clinics’ relationship to a university research hospital helped to ensure that the treatments followed the prevailing standards of the field and were

informed by current research findings. The treatments represent the “standard” for outpatient care in today’s climate of managed care and brief therapies. The clinicians across the clinics are extensively supervised and subject to on-going trainings. Uniform clinical records are maintained across the clinics in a centralized computer database, as well.

The standard treatment is as follows. At an initial appointment, a clinician (psychologist, nurse, or resident psychiatrist) conducts a 30-50 minute individual evaluation following conventional guidelines for diagnostic interviewing (e.g., Shea, 1988). Based on this evaluation, the clinician determines a diagnosis and prescribes a treatment plan appropriate to the diagnosis. The individualized plan is generally reviewed at a staff meeting with other supervising clinicians or psychiatrists for approval and revision. The prescribed treatment typically includes a combination of individual psychotherapy, group therapy, and medication. The psychotherapy and group therapy may be psychodynamic, cognitive-behavioral, or relational in modality depending on the diagnosis.

For the majority of clients, treatment at the outpatient clinics is prescribed for 12 weeks of individual psychotherapy sessions with additional aftercare recommended; on average, clients attend about 6 sessions. Outpatients with a dual-diagnosis for alcohol or drug dependence and an Axis I disorder are referred to a specialized program for co-occurring disorders because of the complications raised by an addiction. The affiliated clinic or unit providing this specialized treatment usually involves the patient in a more intensive version of the same combination of individual sessions, group treatment, and medication.

The cost of treatment averages about \$1,000 per client for a series of outpatient appointments for a fully compliant referral. This cost is typically covered by the clients’ medical insurance if they are employed, or state medical assistance if they are unemployed or with minimal income. During our research project, funds from a state grant were available to pay for treatment costs not covered by

insurance or medical assistance. In this way, there was no evaluation or treatment cost to any of the referred men in our study. Additional cost was not a barrier or complaint for the client or for the courts. The men did, however, have to go through the billing office at the mental health clinics and fill out a variety of forms. Some men complained that billing their insurance company or registering for medical assistance could be detrimental. They feared it would identify them as a “psychiatric patient” and hurt their reputation or work status.

## **SAMPLING PROCEDURES**

### **SCREENING AND REFERRAL**

As mentioned above, the subjects for our research project (N=479) were drawn from the men arrested for domestic violence against a female partner, and ordered by the domestic violence court to attend a minimum of 16 weekly counseling sessions at the DACC program. The men who attended the initial program intake and screened positive for mental health problems were eligible for inclusion. Judges could also order a man to obtain mental health evaluation and treatment based on the evidence of the case and regardless of screening results, but their referral added only a small portion of additional men to the sample. Exclusions from the referral sample and the characteristics of the final sample are discussed in “Part IV: Screening and Referral.” Background information from those who screened negative or were excluded for other reasons (e.g., refusal or age) was retained to develop a comparison with the referred men.

The procedures for the mental health screening and sample recruitment were as follows: The DACC court liaison scheduled each man for an appointment at one of four locations for program intake (see Figures 1 and Table 1). At the program intake, the court-referred men were administered a background questionnaire, a mental health screening instrument (i.e., Brief Symptom Inventory),

and an alcohol screening instrument (i.e., Alcohol Dependence Scale) as part of the standard assessment procedures. Our research project on supplemental mental health treatment was also introduced along with consent information and forms to participate in the research. The intake staff sent the questionnaires and screening tests via express mail to the research office for data entry and scoring. The screening results were, in turn, sent back to the staff to be used at the orientation meeting the following week, and to the DACC main office for tracking compliance with subsequent referrals. (Some of these procedures were slightly modified to improve consistency and efficiency, as discussed in “Part III: Formative Evaluation.”)

After completing the other tasks of the orientation meeting, the staff met with each man individually to assign an on-going counseling group and to determine his program fee, while others looked over program materials. Each man was also advised of the screening results at that time. The men who screened positive were asked to remain after the orientation meeting for further instructions. The staff presented these men with a simplified list of instructions for referral, contact information for the mental health clinics, and compliance verification forms to present to the clinic (see Figure 2).

The referred men were to call the mental health clinic within a few days and schedule an appointment for a mental health evaluation. They were directed to an administrator, who was familiar with our project and involved in its planning, for a brief screening over the phone and scheduling for clinical evaluation. At the evaluation appointment, the men presented forms to the clinic intake staff explaining the referral and requesting verification for compliance. The clinic was to fax a form indicating the diagnosis, prescribed treatment, and date of the evaluation to the DACC main office. There was also a form to fax later indicating treatment compliance. (Eventually, a case-manager began to determine compliance by calling the clinic administrator and getting her verification based on the clinic’s computerized records.)

Under the mandatory referral stage of the research project, the DACC office manager would inform the court liaison about referral compliance, along with DACC attendance, due for court review. The liaison presented that information to the court, and the court in turn was to sanction men who were non-compliant with jailing, fines, or extended counseling. The referral process, as it turns out, relied on a long chain of communication. Information needed to be shared between the DACC office and the program intake staff regarding screening results, the staff and the men about referral procedures, the clinical and batterer program to verify compliance, the batterer program and the court liaisons with compliance information, and the court liaisons and the judges regarding possible sanctions.

## **SCREENING INSTRUMENTS**

The mental health screening was intended to identify cases with a high likelihood of having a diagnosable disorder and current symptomatology. The Brief Symptoms Inventory (BSI) was selected as the principal screening tool because of its focus on symptoms associated with psychological distress in the past seven days. It also is sufficiently short at 53 items to be manageable for a program intake that includes other tasks. The BSI generally took about 15-20 minutes for the men to complete with assistance from the staff reading out loud the items to accommodate men with reading deficiencies. The Alcohol Dependence Scale (ADS) (Skinner & Horn, 1984) was also administered to supplement the BSI, given the prevalence of alcohol abuse among the batterer program participants and its association with re-assault. This 25-item test was selected because of its short length and correspondence with a diagnosable disorder recognized by the DSM-IV and warranting specialized treatment. The widely-used Michigan Alcohol Screening Test (MAST; Selzer, 1971), by contrast, focuses on problem drinking *ever* in the past that suggests



alcoholic tendencies. Interestingly, the ADS did not add many men to those already screening positive on the BSI (see “Part IV: Screening and Referral” regarding referral sample).

The BSI is an adaptation of the longer Symptom Check List-90 that generates three global and nine syndrome scores from separate subscales (Derogatis, 1983, 1993). The BSI has been widely used over the last two decades in clinical settings and more recently with prison inmates (Murray, 2001) and with batterers (e.g., Gavazzi, Julian, & McKenry, 1996). The strongest part of the instrument, according to validation studies, is the Global Severity Index (GSI), which summarizes the level of distress across the syndrome subscales (Peterson, 1989). The Global Severity Index is touted as the “most sensitive” subscale in identifying psychopathology and has a test-retest reliability of .90 (Boulet & Boss, 1991). The construct and discriminate validity of the syndrome subscales are less impressive (Peterson, 1989), and factor analyses suggest a unidimensionality of the BSI (Piersma, Boes, & Reaume, 1994). Following the recommended guidelines for male outpatients, men were considered “positive” on the BSI and warranting a clinical evaluation, if they scored greater than 62 out of the maximum 80 points on the GSI or any two of the other subscales.

## **OUTCOME EVALUATION**

As mentioned at the outset, the major and most complex component of our research project was an outcome evaluation of the supplemental mental health treatment. The effect of the referral to the supplemental mental health treatment (i.e., the intention-to-treat) was tested by comparing batterer program outcomes across the three subsamples of referral implementation: voluntary, transitional, and mandatory referral (the procedures for these subsamples are discussed in “Part III: Formative Evaluation” and their characteristics including sample size appear in the beginning of “Part IV: Screening and Referral”). The effect of actually receiving the supplemental treatment was

examined by comparing the treated men to the untreated men (i.e., dose-response). A 12-month follow-up period following program intake was used to collect outcome data on the men's program completion and further abuse of their partners. The outcome variables and follow-up procedures are discussed below.

## **VARIABLES**

### **Outcome Variables**

The two principal outcome variables for comparing the effectiveness of supplemental mental health treatment are batterer program completion and re-assault of the men's female partner. Program completion is of special interest because of the established relationship with re-assault (Gondolf & Jones, 2001) and the court's consideration of dropout as "non-compliance" subject to further sanctions. The women's report of "re-assault" was used because it remains a relatively objective measure of woman abuse, and is the main reason for domestic violence arrests and batterer program referral.

"Program completion" was relatively easy to determine. The research staff obtained the computerized attendance records from the batterer program's main office and coded them for program completion. Program completion was measured as attending the required 16 sessions without dismissal for unexcused absences or failure to pay assessed fees. Re-arrest for domestic violence, violation of a civil protection order, or reporting to a counseling session drunk or high also resulted in dismissal. As part of the last required session, program participants, furthermore, submitted and read aloud a letter reviewing what they had learned and how they planned to behave.

"Re-assault" was defined as physical abuse of the subject's female partner reported by that partner during phone interviews conducted every three months over the 12-month follow-up period.

It was identified through a series of questions that included an open-ended question about how the “relationship is going,” descriptions of any conflicts and their circumstances, and an inventory using the categories of the Conflict Tactics Scale (Straus, 1979). The nature of battering injuries and medical assistance received for those injuries, and the woman’s response to the abuse, was also obtained as further descriptive information to elaborate the re-assault findings.

More specifically, a re-assault was considered any incident that included one of the tactics on the physical aggression subscale of the Conflict Tactics Scale (i.e., push, shove, grab; slap; hit with a fist, bit, kick; hit with something, attempt to hit with something; choke or burn; threatened with a knife or gun; used a knife or gun; forced sex against will). “Severe” re-assault was also considered and defined as use of any of the so-called “severe” tactics of the Conflict Tactics Scale (i.e., hit with a fist, bit, kick; hit with something, attempt to hit with something; choke or burn; threatened with a knife or gun; used a knife or gun; forced sex against will). Also “threats” were identified using a series of items drawn from scales for non-physical abuse (e.g., Marshall, 1992; Shepard & Campbell, 1992; Straus, 1979; Tolman, 1989). These items included the man’s threatening to hit, attack, or harm the woman; kill any person; take away or harm children; harm other people; and kill or hurt self.

The women’s subjective appraisal of their own safety and well-being was used to supplement the threat and assault variables. These variables offer some indication of the subjective experience of the women beyond the behavior of the men, and are highly correlated with reports of re-assault and independently predictive of re-assault (Gondolf, 2002; Heckert & Gondolf, 2004). The women were asked a global question, “How would you rate the quality of your life overall?” They were also asked to rate how safe they feel at that point, and how likely it is that their partners would hit them in the next three months (using a Likert scale).

## **Other Variables**

Several other variables were assessed at program intake to describe the sample, to test for equivalent subsamples, and to serve as control variables in subsequent analyses. A background questionnaire included close-ended questions about the man's demographics, employment, relationship status, past assault, alcohol and drug use, prior social service and criminal justice contact, and the women's previous help-seeking. These questions encompass variables that have been shown in previous program evaluations to be associated with program outcomes (Daly & Pelowski, 2000; Tolman & Bennett, 1990). The follow-up interviews also included variables about the man's and woman's employment status, the man's alcohol and drug use, partner contact, the man's additional social service contact, and the women's additional help-seeking. These offer situational variables that may mediate outcomes, as suggested in previous evaluations (Jones & Gondolf, 2001).

Drinking and drug use were assessed using men's and women's reports about the kind of substance, frequency of use, and frequency of "drunkenness" or "getting high." In our previous research, we found that these reports were much stronger predictors of outcomes than the results of conventional screening tests or scales (Jones & Gondolf, 2001). The men's social service contact and women's help-seeking are based on inventories developed in our previous research (Gondolf & Fisher, 1988; Jones & Gondolf, 2001). The responses are grouped into categories such as alcohol treatment, other counseling, and battered women's services; and used as categorical ("yes" or "no" response) or interval (total number of services contacted) variables in the analysis.

Results of the BSI subscales were also used to measure the nature and extent of mental health problems. A positive score on each subscale helped to indicate the nature of the problems, and the total number of "positive" subscales and a maximum score on any subscale were used to represent

the extent or severity of the problems. These BSI variables were used to predict compliance and outcomes alone and in combination with other background characteristics.

## **FOLLOW-UP**

The female partners were interviewed at program intake and 3, 6, 9, and 12 months after program intake primarily to obtain reports of re-assault. The follow-up interviews ask a combination of open-ended and closed-ended questions about the women's relationship status, abusive behavior and circumstances, and help-seeking and additional intervention. The interview was previously developed to encourage the women to develop their "story" (Heckert, Matula, & Gondolf, 2000). This procedure appears to increase rapport and disclosure of abuse, according to debriefing questions about the interviews at the conclusion of the previous study (Gondolf, 2000a), and takes from 10 to 20 minutes to administer. Some interviews lasted 45 minutes if the woman had suffered severe abuse.

The follow-up intervals are selected for practical and conceptual reasons. The 3-month follow-up is soon enough after program intake to minimize loss of the women due to change of residence or telephone. The 6-month interview helps to further build rapport and provide tracking information for these women. Interviews at this time period also capture the majority of first-time re-assaults. We found in our multi-site evaluation that over two-thirds of the first-time re-assaults occurred within six months of program intake and continued to decline after that (Gondolf, 1997a, 2000c). The 9-month interview appears essential in maintaining contact with the very transient women for the full follow-up period. In our previous multi-site evaluation, the 12-month follow-up captured over 80% of first-time re-assault committed in a 2½-year period after program involvement

(Gondolf, 2000c). It also offers a cumulative period that is comparable to the majority of program evaluations in the field.

The women were tracked using the contact forms of names and addresses obtained from the male subjects at program intake and updated with each follow-up interview. They were informed about the study, given detailed information about safety measures, and asked for consent over the phone (as done in our previous multi-site study). We acknowledge that safety and ethical issues are involved in conducting follow-up interviews with a clinical sample of battered women and have taken care to develop procedures for addressing this. We report these procedures in a recently published book chapter (Gondolf, 2000a), along with evidence of the minimal negative impact of the interviews. The women were paid \$10 for each completed interview through a check mailed to an address they designated. We have found the payment contributes to the response rate, as well as offers a suitable compensation for the interviews.

Our tracking and interview procedures have previously achieved an unusually high response rate for a resistant clinical sample of this kind (Gondolf, 1997a, 2000c). In the current research project, the response rate for the full 12-months was 65% for the female partners; 79% of the woman completed at least one interview during the follow-up. Women who could not be reached by telephone were mailed a letter at the address they offered at their initial interview at the time of the man's program intake. In the letter, they were asked to call the "program evaluation" office for an interview, or to fill out an enclosed questionnaire and return it by mail. Approximately, 6-9% of the total respondents chose to complete and mail the questionnaire at any one follow-up interval. Any new partners identified by the men or by their initial female partners were also interviewed, and any reports of re-assault or abuse from these women were combined with the initial woman's reports (44 new partners were identified). In our previous multi-site study, we also interviewed new female

partners and found their reports did not increase the re-assault rates based solely on the initial women's reports (Gondolf, 1997a).

## **ANALYSIS**

The analysis examined the effect of supplemental mental health treatment on the follow-up outcomes using both intention-to-treat and dose-response analytical frameworks (Efron & Feldman, 1991). The steps in this analyses included testing for selection bias in our subsamples, cross tabulations of the outcomes by the referral and treatment subsamples, and confirmatory logistic regressions controlling for background variables. The lack of any substantial effects in our cross-tabulations precluded the utility of computing effect size, and the low referral compliance limits the robustness of more complex analyses, such as propensity score analyses.

As a first step, we tested for selection bias across our subsamples representing the referral stages and for non-response or attrition bias in the 12-month follow-up of female-partner interviews. The subjects, and partners of the subjects, comprising the referral subsamples were compared on background variables and BSI results, measured at program intake, to assess whether the referral implementation produced equivalent subsamples. We also examined the possibility of non-response or attrition bias in a similar fashion.

To determine whether the supplemental mental health treatment reduced the likelihood of program completion or re-assault, we cross-tabulated the outcome variables by the subsamples for referral stage and treatment compliance. Cumulative variables were used for re-assault, severe re-assault, and threats. (At least one of these tactics reported over the 12-month follow-up was coded as "1", and no tactics as a "0".) Only those subjects with an interviewed partner for the full 12-months were included in the analysis (n=308), but subjects with an interview any time during the follow-up

(n=479) produced nearly the same results. Chi Square statistics or Fisher's exact test (for the two-way treatment tabulations) were used to assess whether the outcomes were significantly lower for the men under mandatory referral or who actually received treatment. However, these significance tests are not likely to be very instructive in our research project because of their sensitivity to sample and cell size (e.g., the compliant men receiving treatment is small) and their attempt to generalize from probability sampling to a larger population (e.g., our referral sample is a very purposive one) (Cohen, 1994).

To control for the possibility that the cross-tabulation results were influenced by differences in subsample characteristics, we computed multivariate confirmatory analyses using logistic regressions to predict the outcomes. Blocks of variables for demographics, relationship status, past behavior, prior domestic violence, and past intervention were entered using forward stepwise procedures within each grouping. (The principal background variables and block categorization appear in "Part IV: Screening and Referral" on Table 5 listing the characteristics of the batterer program participants.) Where several measures were available for a particular variable, such as alcohol use, the measures were rotated into the logistic equation to determine the most influential one. A variable indicating the intention-to-treat condition was then directly entered into the equation in order to examine its significance in predicting the cumulative re-assault over the 12-month follow-up or program completion. That variable was the designation of the referral stage with "mandatory referral" representing the "intention-to-treat." Similar regressions were computed for the dose-response condition by first entering the variable for "obtaining an evaluation" and then computing a regression entering "receiving treatment." A category for "missing" was used with these two variables to account for those cases that were under the voluntary and transitional stage, and thus allowed for a larger sample and the statistical power it affords the analysis.



## ENDNOTE

1. The longer nine-month batterer program in the multi-site evaluation did show a significantly lower rate of severe re-assault and repeat re-assault; however, the women's "quality of life" was also significantly lower for the nine-month program and the other outcome indicators were not significantly different (see Gondolf, 1999c). The long program with men of higher socio-economic status was also not a significant predictor for less severe re-assault when controlling for batterer characteristics across the sites ( $p=.65$ ). Additionally, we found some evidence that the lower severity of re-assault may have been related to the alcohol treatment provided by the longer program rather than simply its longer duration (Jones & Gondolf, 2001). Our conclusion was, therefore, that the program length in itself did not substantially affect program outcome. Other factors such as delay of program enrollment, low rates of program completion, acquisition of additional services and treatment, and inconsistent court sanctions for non-compliance may counter the effect of program duration.

## **PART III: FORMATIVE EVALUATION**

### **INTRODUCTION AND METHOD**

We report first on a formative evaluation of the implementation of the supplemental mental health treatment for batterer program participants. This evaluation includes a description of the issues and problems that emerged in coordinating the courts, batterer program, and mental health clinic involved in the research project. This formative evaluation formally began within the first year of the project as challenges to the implementation emerged. We systematically documented those challenges based on our observations of the agencies involved in the study and weekly conversations with administrative staff at these agencies. The documentation was necessary given the threats to our initial research design which we were forced to eventually revise. After revising the design (as discussed below), we also outlined a plan for monitoring and documenting the implementation of the new design. These steps added further information to our formative evaluation.

It became readily apparent that much of what we had to learn would be from the implementation process, and a formative evaluation would, no doubt, lead to information that would help in interpreting our outcome evaluation—that is, the effectiveness of supplemental mental health treatment in improving batterer program outcomes. We have previously experienced what critics to narrowly-focused experimental evaluations have argued—namely, that program or treatment context can substantially influence outcomes (see Dobash & Dobash, 2000; Guba & Lincoln, 1989; Pawson & Tilly, 1997). Our previous experimental clinical trial of culturally-focused batterer counseling produced outcomes that are likely to have been neutralized by the relationship of the experimental group to the batterer program and the courts (Gondolf, 2007). Therefore, a formative evaluation seemed an important addition to the larger research objective.

Our formative evaluation drew on direct observation of the agency procedures, participation in training and supervision meetings, debriefing interviews with administrators, and informal comments from staff and clients. The principal investigator and a research assistant compiled fieldnotes on all of these aspects; separately summarized the main issues, challenges and themes in the notes; and discussed and compared our summaries to help verify our conclusions. These conclusions were further clarified, focused, and validated through the latter two of three advisory committee meetings. The committee was comprised of representatives from each of the collaborating agencies and three researchers from the mental health field specializing in the treatment of violent psychiatric patients.

In terms of observations, there were two trainings of batterer program staff, periodic observations of the screening at program intake, and visits to the domestic violence court with the batterer program liaison. During the mandatory referral stage, we met four or five times with administrators at the two mental health clinics in order to review the referral and reporting procedures and discuss feedback from the other agencies. The principal investigator and research assistant conducted joint interviews with clinic administrators, batterer program staff, and one court representative. We were also in regular contact with participating staff at each agency about the protocol. Our case-management exposed lapses in the notification reporting procedures that were, in turn, discussed in person or by phone with various agency staff. Examples include positively screened men not being notified of their required referral, and clinic verification of a man's obtaining a mental health evaluation not being sent to the batterer program.

In a sense, the implementation of the mandatory referral represented a kind of "system audit" that information about the procedures to identify inconsistencies and breakdowns (see Pence & McDonnell, 1999). The information was then used as feedback to various agency staff to help correct

or adjust their practices. It also exposed to the research team other remedies to improve the implementation, such as additional training, meetings, or procedures. In this way the formative evaluation was very much participatory or “action” research. The information being collected was used to modify what was being observed.

Finally, our analysis was based largely on the referral protocol and procedures that were developed over the course of about six months in consultation principally with batterer program and court administrators (see Table 1). The information we collected was contrasted with these protocol and procedures to expose inconsistencies and breakdowns, and generate possible modifications to the referral “system.” The protocols and procedures remained relatively intact with minor changes in the means of reporting (e.g., sending email verification of compliance to the case-manager rather than faxing forms to the batterer program office staff) and increased monitoring of referral notification following screening.

### **DESIGN MODIFICATION**

The research design was initially to be an experimental evaluation of supplemental mental health treatment. Men who screened positive at batterer program intake were to be randomly assigned to mental health treatment, and the outcomes of the experimental group (i.e., treated men) and control group (i.e., non-treated men) were to be compared. We had success previously in conducting two experimental evaluations at the research site (Gondolf, 1997b; 2007), but these evaluations were conducted within the batterer program without outside referral. We also had several preparation meetings with administrators with the collaborating agencies. They were exceptionally supportive, interested, and eager to participate in the research project and accepted its design. Their personal observations, experiences, or expertise appeared to contribute to this enthusiasm, and several

saw benefits for their component as a result of the project. For instance, court clerks felt that the protocol of the mandated referral would assist their record keeping, and the judicial administrator anticipated it giving the court more options for difficult cases.

However, several unexpected events interrupted the intended research design and forced a modification of random assignment. First, the district attorney overseeing the domestic violence court objected to the random assignment at our initial start-up meeting. A highly publicized murder had recently occurred in which a convicted domestic violence perpetrator entered a large municipal church to kill his estranged wife. The district attorney was concerned about the possible public reaction, court liability, and ethical issues associated with identifying high risk men through screening and then not doing anything about it. He also foresaw objections to some men being voluntarily referred to mental health treatment and others having the extra burden of being mandated, and the difficulty in keeping track of which was which. The district attorney did accept, however, the idea of an initial voluntary screening phase progressing into a court-mandated one.

A second major disruption was an administrative scandal at the batterer program. During the month in which the research project was to begin, the assistant director disclosed that the director had been embezzling agency funds and also committing health insurance fraud. The financial viability and ultimately the continuation of the agency were thrown in question as a result. As the agency board investigated the extent of the problem and its impact on the agency, we began to screen and refer men on a voluntary basis for supplemental mental health treatment. This step provided an opportunity to develop and test the screening procedures and also to develop a group of voluntary referrals to compare to mandatory referral. As it turned out, less than 5% of these “voluntary” men complied with treatment and conveniently created an “untreated” comparison group.

Thirdly, the institutional review board (IRB) associated with the mental health clinics determined that an “experimental” study posed special human subjects concerns. Our contention, and that of our university IRB, was that the referrals would be under the usual “standard of care” and not subjected to an experimental treatment that was new, innovated, and untested. The clinic IRB also decided that it was unethical to identify men warranting mental health treatment and then to withhold that treatment from them. They must be given the opportunity to seek treatment on a voluntary basis—that is, be notified about the screening results and presented with referral information. These IRB decisions reinforced using voluntary referrals as a comparison group to a later mandatory referral group, rather than a randomized assignment of treatment versus no treatment.

The nature of these events is admittedly extreme and unusual. They do, however, reinforce the concern about “real world” circumstances in program evaluation, and the difficulties in conducting experimental designs. As we experienced in this current research project, the resistance seems to increase or be more likely when the random assignment involves a variety of agencies, and when withholding treatment is one of the conditions. In our previous clinical trials, we were able to control the random assignment as part of the agency discretion (i.e., assignment to its various batterer counseling groups) and offer some alternative “treatments” as a control group (Gondolf 1997b; 2007). Our comparative group design (i.e., voluntary referrals compared to a latter group of mandatory referrals) did offer, as it turns out, an opportunity to study the array of issues, barriers, and challenges to implementing mandatory referral and the coordination that must accompany it. The modified design also helped to expose naiveté regarding “coordinated community response,” or at least coordination of the domestic violence courts, batterer programs, and mental health clinic. Our previous experience with coordinating various referrals and collaborations, as well as conducting

numerous studies at the current site, led us to believe that implementation of the protocols and procedures would be relatively straightforward. That was not to be the case.

## **IMPLEMENTATION PROBLEMS**

We encountered a variety of inconsistencies and breakdowns of the referral protocols and procedures as we attempted to administer the mandatory referral. There were problems in nearly every step of the process—from the initial screening at program intake to court sanctions for non-compliance (see Table 2). Following the initial training with batterer program staff, the four staff conducting program intake began to vary their screening approach. Some were more authoritative than others; others wanted to accommodate the men's questions and resistance to screening and research follow-up. Also, the number of men appearing for the weekly intakes at four different locations fluctuated, and changed the level of interaction between the staff and program participants. The staff also questioned the utility of the screening and other paperwork, such as the background questionnaire and consent forms, and they resented the burden which screening placed on them. They had other tasks associated with program intake, such as determining the sliding fee for each man and assigning him to an appropriately located counseling group.

The men who screened positive on the BSI administered at program intake were to be notified of the test results and the referral procedures at the orientation meeting scheduled for the following week. The notification was initially inconsistent because of men missing the orientation session and going directly to an on-going program group. Also, if a staff member was absent or the meeting room was occupied, the intake-orientation sequence was thrown out of order and several men would miss their screening or notification. Furthermore, intake staff occasionally delayed submitting the

completed BSI materials to the main office for scoring, and consequently no test results were available for the orientation meeting.

The caseload and priorities at the mental health clinic contributed to some lapses in communication and linkage. Some referred men complained about the difficulty in contacting the clinic for an initial appointment, or about problems with clinic billing procedures (evaluation and treatment costs were to be covered through various insurance, welfare, and compensation plans). The clinic also did not always fax the forms verifying compliance to the batterer program. The men lost the forms, the clinic staff forgot to fill them out, or the man was not identified as needing verification. Additionally, the referred men were often uncooperative and resistant to the evaluation and difficult to diagnose as a result. Such men were not recommended for treatment but considered compliant to the referral for an evaluation.

The final area of problems was with the courts. For one, the batterer program office secretary was to communicate the referral compliance, along with batterer program attendance, to the court liaison. This information was sometimes not available or the liaison was absent or failed to collect it. One liaison stopped presenting the compliance information to the court since the judges did not necessarily respond to it. Another occasionally withheld the information if he thought the man was otherwise in good standing with the batterer program. Various district attorneys would sometimes waive non-compliance to move cases along and resolve outstanding charges. The main and persistent challenge was to the judges' inconsistent response to the referral non-compliance. The leverage or "accountability" of the mandatory referrals was thus weakened. What amounted to the lack of sanctions for non-compliance was the result of two related circumstances. As the mandatory referral began, the city consolidated its judges (actually "district magistrates") as a cost cutting measure, and laid-off the judicial administrator who coordinated the domestic violence and other specialty courts.



The number of judges rotating through the domestic violence court, therefore, expanded from five to over ten. The new judges were less familiar with domestic violence issues and exercised more discretion in case decisions than the previous experienced and supervised judges.

In sum, the complexity of the communications required for the mandatory referral made it vulnerable to inconsistency and breakdown. The number and diversity of frontline staff across different agencies also appeared to add to the problems. The supervision of the staff was initially somewhat lax, and our coordination of the collaborating agencies also sporadic. We had inadvertently assumed that the independent agencies would manage the referral process and adjust as need be. We also assumed that our contacts primarily with the agency administrators would be sufficient to implement the mandatory referral and maintain consistent procedures among their respective staffs. The coordination and communication required for the mandatory referral clearly warranted more than establishing protocols and procedures.

## UNDERLYING ISSUES

It is difficult to construct a causal model for the implementation problems. Several issues are likely to be interrelated and variable. We were able, at least, to identify organizational and structural issues that appeared to contribute to the inconsistencies and breakdowns discussed above.

Interestingly they resonate with other observations of efforts to develop “coordinated community response” to domestic violence cases (e.g., Gondolf, in press-a; Visher, Newmark, & Harrell, 2006).

Of particular note were administrative absenteeism and turnover, administrative-staff gaps, client overload, and differing agency priorities.

Administrative turnover was a major issue since our screening and referral relied on key staff at each collaborating agency. The “key staff” were administrators who helped develop the referral

procedures and protocol and served as the representative of their respective agencies. When they were absent from their position or left the agency for other work, breakdowns were more likely. Early in the implementation of the mandatory referral, the judicial administrator was laid-off under city cost-cutting, and the batterer program director was dismissed for financial mismanagement and fraud. The judicial administrator was not replaced, but existing staff did assume the program director position. Also, the district attorney overseeing the specialty courts was promoted and a replacement from a different court was appointed, and the clinic director was absent for a couple of months because of surgery and later given additional responsibilities that stretched her further.

Fortunately, the new batterer program director replaced the former assistant director and office staff with other program employees, and maintained the continuity of the services. We were heavily involved in the transition to help and observe the program operation for our research project. From what we could observe at the program, in staff interviews, and follow-up interviews, there was no interruption or consequence to the batterer program sessions or staff performance during the administrative overhaul. The administrative changes did, however, initially weaken the attention and supervision of the screening and referral process, as did the absence or departure of other agency administrators.

Related to the administrative turnover was the gap between administrators and the frontline staff enacting the screening and referral procedures. We relied heavily on the knowledge and commitment of the administrators to implement the mandatory referral at their respective agencies. For a variety of practical reasons, they were not always able to do so. The demands of their respective positions, the need to maintain the existing services, and the turnover and absenteeism discussed above often precluded direct and continuous supervision of the screening and referral procedures. The staff resistance, barriers, and misunderstandings were, therefore, not fully realized

and addressed at times. We also observed at the batterer program so-called “training decay” as the research project progressed. Administrators and staff began to forget, misrepresent, or misunderstand some of the referral procedures as they were initially taught.

The Judicial Oversight Demonstration (JOD) project to enhance community coordination identified a consequence of such projects (Visher, Newmark, & Harrell, 2006). The increased coordination, referral, and accountability are likely to further overload the collaborating agencies. More cases or clients are to receive services, monitoring, and management, and staff are required to do additional tasks and assume more responsibilities. Even with additional funding the overload often persists and contributes to poor implementation. The batterer program in our research project initially obtained outside funding for a case-manager or system-coordinator. Similar to some agencies in the JOD project, those funds were diverted to other needs within the agency, and the manager and coordinator responsibilities were absorbed by existing office staff. Existing staff assumed the responsibilities of the proposed positions in previous projects at this site (see Gondolf, 2007, in press-a).

Another aspect of overload is the demands of the usual client or case load. The courts, clinic, and batterer program in our study were all operating beyond capacity prior to our research project, and each had experienced recent funding cutbacks. They were insufficiently staffed, supported, and equipped (e.g., inadequate computers or data systems). The mental health clinic, for instance, received approximately 100 phone calls per day, and clinicians were evaluating as many as 15 clients per day. The occasional difficulties in contacting the clinic or obtaining verification were understandable in this light. Similarly, the batterer program had intakes for 80-100 men a month and 20 on-going groups to manage. The program need to collect fees, record attendance, track

compliance, and coordinate with the court and probation precluded attention to mental health referrals.

Lastly, the agencies involved in our research project differed in orientation and priorities. There was no outright conflict or competition, but there was sometimes a clash of purposes, assumptions, or expectations. Probably the most obvious in this regard was the difference in the punitive orientation of the court and the accommodating approach of the clinic. The court expected and required the men to obtain a mental health evaluation, and assumed that coercion from possible sanctions was sufficient to have them complete that task. As the clinic staff explained to us, the clinicians tended to rely on their clients' wanting help or treatment. Their clients are typically motivated to present and discuss their mental health problems and needs. The clinicians, moreover, reported a reluctance to be involved in court-mandated cases because of the time, persuasion, and documentation they require.

### **RECOMMENDED REMEDIES**

The underlying issues expose the need, not only for further staff support and supervision, but also for some structural changes or reorganization. Our immediate response to the identified problems was to take a more active role in the referral implementation. We discussed our observations among our research staff and advisory committee members, and initiated intermediate corrections. During what became a transitional referral stage between voluntary and mandatory referral, we recommended or reinforced closing one of the four intake sites, contacted the intake staff by phone or email to confirm referral notification, adjusted the intake schedule to ensure the sequence, and redirected clinic verification to the research office via e-mail. The clinic administrator began to check referral compliance through the clinic's database, designated experienced clinicians to

conduct the evaluations, and encouraged motivational interviewing with resistant referrals. The referred men were also advised to make sure they identified themselves by the batterer program name (DACC), so the clinic would know to promptly schedule an appointment.

One more extensive remedy would have been additional on-going staff training and supervision. We retrained the intake staff at the batterer program once during the research project, and needed to do that every few months, perhaps in briefer meetings, to sustain interest and consistency. Part of the reason why this did not happen was because of the additional cost of paying the contracted staff for their participation in training, and the difficulty in scheduling all the staff for a joint training time. Most of them had other jobs and families to consider, along with leading two batterer groups a week. The principal investigator or research assistant did observe the program intake every 2-3 months, and the domestic violence court occasionally. When we did, we were usually able to assist the intake staff or court liaison in completing their tasks, as well as, question or reinforce referral procedures. More regular supervision and oversight would have, no doubt, been beneficial.

The staff retraining and supervision would help to alleviate two other issues—the consequences of administrative turnover and the administrator-staff gap. In a sense, additional training and supervision would augment the administrative oversight when it was interrupted and incomplete, as well as stimulate communication within the agency ranks. Our involvement in the implementation did this to a degree, at least enough to establish consistency through all but the court response to non-compliance. When this level of consistency was obtained, we considered the referral to be in a mandated rather than transitional stage. (The court inconsistencies were never fully resolved. The issues raised by the new judges and their discretion, as well as the district attorneys need to resolve some cases, were beyond our resources and timing.)

The major remedy is an organizational one. The remedies discussed above ultimately need to be sustained and routinized through a structural modification, otherwise they are likely to be undone through the reoccurrence of the problems we initially encountered, such as administrative turnover. We established an independent case-manager and system-coordinator for this purpose, and, as is seen in the compliance results discussed in “Part IV: Screening and Referral,” this organizational addition had a substantial impact. This person was one of our research assistants with previous program experience, administrative training, and familiarity with our research project. She attempted to call each referred man within a few weeks of batterer program intake to reinforce his compliance, help negotiate referral problems, and collect information about referral experience. She also was to maintain systematic contact with the agencies and facilitate the communications between them. For instance, she sought and received a list of the men notified about referral at the program orientation session, and the names of men who complied with the mental health referral. She also met periodically with administrators to discuss any issues or concerns, as well as to convey the perspectives of the other collaborating agencies.

The fact that she was under the direction of the research director, rather than a particular agency, helped in two ways. One, she was privy to information coming from the research and various agencies to identify problems and ways to address them, and, two, she was not consumed by other agency demands, biases, or directives. She had the freedom to talk with the batterer program participants in her case-manager role and also converse with administrators and staff in her system-coordinator role. Consequently, she had a fuller picture and more leverage.

This remedy was perhaps an exceptional one in that the case-manager and system-coordinator position was part of the research project and under the supervision of the principal investigator. The question beyond the research project is where such a position would best be situated to maintain

relative independence and yet have the support and authority to affect cooperation and correction. Also, identifying a person with the necessary combination of skills and experience seems essential. Most available people are likely to have background in one agency or another, and the orientation and biases to go with that. Perhaps an advisory committee of some kind could help offer the checks-and-balances that the principal investigator and research team were able to provide in our project.

There is at least one community operating with this approach. Spokane, Washington maintains a coordination office for domestic violence intervention. Local services pay annual dues to fund the office and its staff. The coordinator operates independently of the separate agencies while conducting many of the tasks discussed above. These tasks include training meetings and conferences, as well as system monitoring and auditing, and negotiating with administrators for procedural changes and adjustments. This structure goes beyond the domestic violence councils in many committees that often encounter the embedded problems we did, and need either strong leadership or outside initiatives to solve them (Allen, 2006).

An additional structure remedy may be integrating batterer programming and mental health treatment within the same agency. Integration would reduce the complexity of the referral process and reduce the communication breakdowns. Consolidating staff and services would also likely improve consistency and accountability. As discussed further in “Part VI: Summary and Discussion,” we are aware of at least two batterer programs that are embedded in so-called behavioral health units of hospitals. At the one in Calgary, Canada, men are initially evaluated at the batterer program attached to a women’s center. Men warranting mental health treatment are sent to the behavioral health unit for combined domestic violence education and mental health treatment instead of attending the conventional batterer program. A hospital-based batterer program in Fond du Lac, Wisconsin, conducts an extensive evaluation of court-referred batterers and provides supplemental

treatments within the hospital setting along with the batterer educational sessions. Conversely, the AMEND batterer program in Denver offers men individual mental health treatment in addition to the batterer program, as well as special attention in the batterer group sessions. The main objection to such integration is the extra layer of staffing and cost it requires, in comparison to referring to existing agencies and services. There is some concern, too, that the integration may diffuse the needed focus on stopping domestic violence and convey the impression that the violence is really a mental health problem.

### **SUMMARY**

Several unexpected events precluded our initial experimental design and led to sequential comparative samples of voluntary and mandatory referral. More typical inconsistencies and breakdowns in the referral protocol and procedures prompted more direct involvement of the research team in implementation of the project. Besides offering intermediate corrections, we also established a case-manager and system-coordinator to systematically monitor the referral process and ensure its operation. One of the main lessons from all of this is screening and referral across agencies requires more than protocols and procedures. Ultimately some organizational and structural changes appear to be essential in establishing consistency and accountability. There may also be benefits to simplifying the referral process by consolidating services or at least reducing the different steps, agencies, and staff. The implications for our outcome evaluation are two-fold. The involvement of the research team in the implementation of the referral, and establishment of an independent case-manager and system-coordinator, are not typical of what might be done beyond the research project. Also, the intended referral process was, even with this involvement, not as fully implemented as planned. We do not, therefore, know exactly what would be the outcome under “real world”



circumstances, nor do we know the outcome under ideal conditions (i.e., full implementation of the referral protocol and procedures).

## **PART IV: SCREENING AND REFERRAL**

### **INTRODUCTION**

This part of our report presents a series of primarily quantitative studies addressing the utility of screening and referring batterer program participants for mental health problems. The studies together might be considered a “service delivery” evaluation in that they examine the results of screening and referral procedures. They consider the extent and characteristics of the referral sample, the validity and reliability of the screening, and the extent of referral compliance and reasons for it.

The first section below describes the recruitment of our sample of batterer program participants and the characteristics of the sample of men who screened positive for mental health problems. The second section reviews several reliability and validity tests of the screening criteria. The aim here is to examine the possibility of over-screening and alternative screening criteria. The third section considers the men’s response to their referral to mental health treatment in terms of their attitudes and actual compliance.

### **SAMPLE CHARACTERISTICS**

#### **INTRODUCTION**

Our research project is based on a sample of domestic violence perpetrators who appeared to warrant further mental health evaluation and possible treatment based on a variety of criteria. In this section, we review in detail the results of our screening and recruitment procedures along with the characteristics of the resultant “referral sample.” Nearly half of the men screened at batterer program intake were identified as eligible for referral to a mental health clinic. The main source of the eligibility was a “positive” scoring on the Brief Symptom Inventory (BSI). An alcohol screening test

and judge referrals added only a small portion of men. The BSI scorings, however, revealed a variety of profiles and at least a third of the men had maximum scores on one or more of the subscales.

The referral implementation, as discussed previously, resulted in three distinct stages approximating an “intention-to-treat” continuum: voluntary, transitional, and mandatory referral stages. The batterer characteristics across these three subgroups are very similar. The overall population of the batterer program participants, moreover, is similar in characteristics to a 1995 sample at the same site, but with less “stakes in conformity” than program participants at other sites in our previous studies. The “referral sample” drawn from this “intake” population was more likely to be at risk for re-assault, as suggested by more extensive histories of abuse and violence, and of alcohol and drug use. These risk markers not only indicate the appropriateness of more extensive intervention, but the “referred” men were also more likely to indicate that they would benefit from additional treatment.

## **RECRUITMENT RESULTS**

Our sample of batterer program participants who were eligible for referral to mental health clinics was recruited from the Domestic Abuse Counseling Center (DACC) in Pittsburgh, Pennsylvania (see Table 3). The “referral sample” was drawn from the 1043 male domestic violence perpetrators mandated by the domestic violence courts to DACC between June 2004 and April 2006. The total sample was 479 or 46% of the men initially screened at batterer program intake. This includes men who screened positive on the BSI or ADS minus those ineligible as a result of refusing to participate in the study, assaulting someone other than a female partner, being previously admitted to the study, being under age (less than 18 years old), or receiving an invalid score on the BSI. Nearly 13% of the men at program intake were ineligible for one of these reasons. A comparison of

demographic and behavioral characteristics, however, shows no significant differences between the included and deleted cases that might bias the final sample.

The progression of inclusion and exclusion (i.e., being deleted) from the final sample was as follows. Of the total men screened at program intake (N=1043), 13% (n=131) were deleted because of refusal to participate (n=12), abusing someone other than a female partner (n=34), and receiving an invalid score on the BSI (n=85). Of the remaining 912 men, over half (n=515) screened positive on the BSI or ADS. Another 14% (n=73) of these men were deleted for similar reasons: refusing to continue in the study (n=4), the court incident being abuse of someone other than a female partner (n=46), having returned to court and being previously admitted into the study (n=7), and being underage (n=4), not being properly notified about referral according to program records (n=8), or undetermined reasons (n=4). These deletions left a total of 442 men who screened positive for referral and were eligible for the study. An additional 37 men were mandated by a judge for mental health evaluation and treatment regardless of not receiving a positive BSI score for a final referral sample of 479.

## **THE REFERRAL SAMPLE**

The main screening criteria was a positive score on the BSI (Global Severity Index >62 or any 2 of 9 subscales >62), but there were two other criteria for referral. As discussed in “Part II: Research Methods,” the Alcohol Dependence Scale (ADS) was administered and a positive score on that test (ADS >13) led to a referral as well. However, only 15 men were eligible for referral based on this test alone. Another 31 men who screened positive on the ADS also had positive BSI scores. Also, judges maintained their authority to refer men to a mental health clinic for evaluation and treatment based on information presented at the man’s hearing. Judges, as discussed previously, were

very inconsistent in their referrals and monitoring referral compliance. Judges added an additional 37 men to those already screened positive on the BSI or ADS, and 36 men who had already screened positive for a referral. The total of 73 judge mandated men accounted for 15% of the 479 men who were eligible for referral. In other words, only 73 men instead of 479 would have been referred without the screening process. Our court observations suggest however that some of the judge referrals were made because the judges were aware that our screening system and monitoring of compliance was in place, as well as coverage of any clinic costs. So the actual number of men referred by judges alone would likely be much lower under normal circumstances.

## **BSI SCORINGS**

The batterer program participants who screened positive on the BSI and entered the “referral sample” were identified in a variety of scorings on the BSI (see Table 4). Approximately 60% of the “positive” men scored above the cutscore (>62 of possible 80) on the Global Severity Index (GSI). The remaining 40% qualified by scoring positive on at least two of the subscales. There was a wide distribution across the subscales with no prevailing profile. Interestingly, men were most likely to score positive on the depression (59%), paranoid ideation (72%), and psychoticism (57%) subscales. The latter two subscales suggest hostility and suspiciousness, and alienation and withdrawal. A substantial portion of the men showed an extensive array of symptoms, and severe psychological distress generally associated with such symptoms. Specifically, a third of the men scored the maximum 80 on at least one of the subscales, and a third of the men scored positive on 6 or more of the 9 subscales.

The Alcohol Dependency Scale (ADS) identified only 9% of the men as having symptoms of alcohol dependency, and most of these men also scored positive on one of the BSI subscales. Our

previous use of the MAST (Michigan Alcohol Screening Test, Selzer, 1971) at the same site showed over half of the men (56%) screening positive for alcohol problems and possible alcoholism; and the alcohol dependency subscale of the MCMI-III (Millon Clinical Multiaxial Inventory; Millon, 1994) suggested that approximately 30% of the men were alcohol dependent. (The MCMI is a 175-item test with 24 subscales that correspond to Axis II (personality disorders, e.g., antisocial or narcissistic personality) and Axis I (major disorders, e.g., major depression, schizophrenia) diagnoses, and 3 modifying indices used to adjust subscale scores for excessive disclosure, desirability or debasement.) The ADS strictly measures symptoms of active alcohol dependence, whereas the MAST identifies problem behaviors associated with heavy alcohol use ever in the past.

## **REFERRAL STAGES**

As the screening and referral developed, we found three distinct referral stages reflected in the implementation process, and as result the formation of three comparative subsamples of men: 1) the initial voluntary referral phase (June 2004-January 2005; n=182), 2) a transition phase with inconsistent mandated referral (February 2005-August 2005; n=166), and a phase of enhanced mandated referral, including oversight by a case-manager and system-coordinator (September 2005-April 2006; n=131). The three stages provided three comparative groups that form the basis of our study of “intention-to-treat” effectiveness. Our original design was to compare a non-referred group to those who were referred with the “intention” that they would receive a mental health evaluation with the likelihood of further treatment. The batterer program participants referred voluntarily to a mental health clinic offered our “non-referred” comparison subsample since less than 7% of these referred men made an appointment for an evaluation, and fewer actually obtained an evaluation or received treatment. Voluntary referral, rather than withholding referral, was used to establish a “non-

referred” subsample because of human subjects concerns. The referral at least gave men with an identified problem the option of receiving further treatment, rather than deny them information that could be of assistance to them.

The transitional referral group was an unintended stage that emerged in the efforts to establish a “mandatory referral” in which the men were required by the court to comply with the mental health referral or receive further sanctions from the court. As discussed in “Part III: Formative Evaluation,” our monitoring and observations of the implementation revealed that, for a variety of reasons, the procedures and protocol of the mandated referrals were not being followed as designed. One main reason was a turnover of administrators at the batterer program, as well as inconsistent conformity to the procedures. These shortcomings were gradually addressed through staff meetings, additional trainings, and new personnel including a case-manager and system-coordinator. The principal investigator and system-coordinator finally deemed the procedures operating as fully as practically possible and being monitored through case-management and system observations. (The one remaining inconsistency was failure of judges to respond to sanctioned men who did not comply with the mental health referral on a regular and consistent basis.) At that point, we considered referred men to be under “mandatory referral” and representing an “intention-to-treat” subsample.

## **SAMPLE COMPARISONS**

The characteristics of the population of men entering the batterer program during our study recruitment (N=1043 both referred and non-referred men) had nearly identical demographics and behavioral characteristics as the sample of men from our 1995 study at our Pittsburgh site (see Gondolf, 1999a). For instance, half (50%) of the men were African-American, only a quarter (27%) had some college education, less than a third (27%) were married; nearly half (45%) were

unemployed, and about half (45%) had been previously arrested for non-domestic violence crimes (see Table 5). In the previous study, however, this site had significantly higher rates of African-American, less educated, unemployed, and un-married men than the program participants at the three other batterer programs in our multi-site evaluation (i.e., Houston, Dallas, and Denver). The prior arrest records were similar across the four sites. In other words, the Pittsburgh site appeared to have less “stakes in conformity” than the other sites and consequently greater risk for re-assault (Sherman, 1992, pp. 171-187).

In our current research project, the sample of men referred to a mental health clinic (N=479) was more likely to report problematic behaviors than those who did not screen positive on the BSI and were not referred (see Table 5). Most notably, they were more likely to have a parent who hit the other, used drugs, been drunk at least monthly, threatened their partner in the past 3 months, severely assaulted their partner in the past, been previously arrested for domestic violence and felt likely to hit again, and received psychological treatment in the past (e.g., psychological treatment: not referred 14% vs. referred 34%;  $p < .001$ ; N=1043). The referred men were also nearly twice as likely to indicate that they would benefit from mental health treatment from “some” to a “great extent” (not referred 16% vs. referred 34%;  $p < .001$ ; N=1043).

The characteristics of the referred men remain relatively the same across the referral implementation or “referral stages.” The demographic and behavioral characteristics are nearly identical across the three subsamples from the referral stages (see Table 6). The only significant differences are with a lower rate of men scoring positive on multiple subscales of the BSI during the mandatory referral stage (voluntary 36% vs. transitional 36% vs. mandatory 25%;  $p < .05$ ; N=479), and having previously attended a batterer program (voluntary 18% vs. transitional 23% vs. mandatory 12%;  $p < .05$ ; N=479). We attempt to control for the influence that these or other variables might have



on a direct comparison of the subsample outcomes (i.e., results of cross-tabulations) through a series of confirmatory logistic regressions.

## **VALIDITY AND RELIABILITY OF BSI SCREENING**

### **INTRODUCTION**

As indicated in the “Sample” section above, nearly 50% of the batterer program participants screened positive on the BSI (Brief Symptom Inventory) administered at program intake. This contrasts with less than 25% showing evidence of a major psychiatric disorder on the MCMI-III in our previous research at this site (Gondolf, 1999b). To examine further the possibility of a high-level of over-screening or “false positives” on the BSI, we conducted reliability and validity tests with the BSI using our population of batterer program participants. These tests might also help identify protocol revisions or alternative screening approaches to reduce the percentage of false positives.<sup>1</sup>

Our reliability test consists of re-testing a subsample of men in our study to see if their BSI scores remain constant over time—that is, are “reliable.” Or, do their scores decrease at the re-test suggesting that the BSI scores derived at program intake are inflated by the duress of police arrest and court action, and the uncertainties of program enrollment. To examine the concurrent validity of the BSI, BSI scores were compared to the results of two other established psychological assessment instruments with slightly different criteria and approaches. Is there substantial agreement among these tests that are all designed to screen for mental health problems? Or does the BSI show a higher rate of positive screens than the other instruments? The validity of the BSI was further examined through tests of predictive validity. That is, we considered the relationship of the overall BSI positive screen and positive BSI subscales with batterer program outcomes, specifically program completion and re-assault of one’s partner during the 12-month follow-up. Our initial assumption was that

positive BSI scores are associated with program dropout and re-assault. If they are not “predictive” of outcome, is there a different cutscore that would increase the correct prediction? Each of these reliability and validity tests are discussed separately below with details of the methods, results, and implications.

## **REFERRAL CHARACTERISTICS AND DIAGNOSES**

The main objective of the BSI screening was to identify batterer program participants who warranted mental health treatment. Two initial considerations of the referral sample (i.e., the men who screened positive on the BSI and were eligible for clinic referral, N=479) posed a contradiction. The “at risk” characteristics of the referral sample appeared to endorse the screening results; however, the clinical diagnoses of the men who did comply with the referral (discussed below) suggest excessive over-screening—that is, a high level of false positives. These observations led us to consider further tests of reliability and validity.

### **Batterer Characteristics**

A comparison of the characteristics of the referred men to the negatively screened men not only appeared to corroborate the screening, but also offered preliminary evidence of concurrent validity. The demographic and past behavior associated with a positive result on the BSI represented risk factors commonly associated with criminal re-offense: a lack of “stake in conformity” and a higher level of prior violence and criminality (see Table 5). The BSI-positive men were also more likely to see themselves benefiting from domestic violence counseling, alcohol abuse treatment, and psychological treatment. On this basis, the BSI appears to be appropriately identifying men at high risk for re-assault, and therefore those men warranting supplemental treatment or intensified

intervention. The characteristics, however, do not necessarily substantiate the presence of mental health issues or their effect on the batterer program outcome.

## **Clinical Diagnosis**

The clinical diagnoses present a kind of “predictive” validity of the BSI screening.<sup>2</sup> That is, they show whether the referred men warrant mental health treatment as the positive screening would suggest. As discussed in “Part II: Research Methods,” the clinical records of the men under our mandatory referral stage were obtained from the clinics, and the men’s diagnosis was identified as the major or first DSM-IV-TR code. One exception was to distinguish the men with a diagnosis of alcohol dependence *and* a second disorder—that is, a “dual diagnosis.” We then tabulated the clinical diagnoses for the men who were referred to a mental health clinic and obtained an evaluation (n=38). The clinical diagnoses for the evaluated men (n=38) suggest much less severe disorders than expected from the BSI results (see Table 7). Only about half of the men received diagnoses that would warrant treatment beyond the batterer program, and most of these were men with alcohol and drug problems. Specifically, over a third (37%) was diagnosed with an adjustment disorder, and most of these men did not receive further treatment. Another third of the men were diagnosed with alcohol or drug disorders, and three-fourths of these men had a dual diagnosis (i.e., also bipolar depression, major depression, or anxiety disorder). A small portion (11%) was diagnosed with a depressive disorder or bipolar depression, and 8% were characterized with an impulse control disorder. The diagnosis of the remaining men (5%) was deferred. Moreover, there was little agreement of the BSI subscales with the clinical diagnosis (see Table 8). Only 11% of those men who scored positive on the BSI depression subscale, for instance, were given a clinical diagnosis related to depression.

## Summary

According to the diagnoses given by the mental health clinicians, the BSI appears to substantially over-screen men for treatment and apparently misdiagnose the men. Unfortunately, this assessment is limited to the very few men who complied with the referral (n=38 of 479). Men who did not obtain a clinical diagnosis may have produced different results. Also, the clinical diagnoses may themselves not be valid given the observations about the evaluations reported in the formative evaluation. The diagnosis for some men was difficult to determine because they refused to cooperate with the evaluation interview. In sum, the BSI appeared to identify men at higher risk for re-assault and warranting referral, but nearly half the men who complied with referral did not receive a diagnosis that warranted mental health treatment.

## BSI RELIABILITY

### Method

To further explore the initial contradictions in the BSI screening, we first considered the test-re-test reliability of the BSI.<sup>3</sup> The re-test was conducted with 98 batterer program participants who completed the BSI at program intake. These men were selected through a convenience sampling of the last 200 batterer program participants recruited for the research project. The BSI was re-administered to 21 men at the end of their case-management interview (2-4 weeks after program intake), 34 men at the conclusion of a batterer program session (6-16 weeks after program intake), and 41 men over the phone as part of the 5-month follow-up interview (8 re-tests were invalid). These three sources for re-test offered alternative screening times and protocols to compare and consider. We also asked the men, after completing the re-test, whether they thought their BSI scores

had changed and why they had changed or stayed the same. The answers to these questions might help us interpret the results of the re-test, as well as expose the men's sense of change.

### **Reliability Results**

The BSI re-tests showed a substantial decrease in positive scorings (indicating a mental health referral) suggesting low reliability. Half (50%) of the men who initially scored positive on the BSI scored negative on the re-test (n=45 of 90). Also, the 8 men who initially scored negative scored positive on the second testing. These men had been ordered by a judge to obtain a psychological evaluation regardless of BSI screening results.

The men's perceptions of whether their scores changed were, moreover, often mistaken. Forty men perceived that their scores had improved; however, the scores of 15 (38%) of those men actually had not significantly changed. The main reasons given for the men's perceived change were a shift in perspective (n=14; e.g., a different outlook), new life circumstances (n=13; e.g., a new relationship, less stress, jail time), the impact of the batterer counseling (n=11), and lowered drug or alcohol use (n=5). Additionally, 17 of the re-tested men believed that their BSI score had not changed at the re-testing. In about half of these cases, the scores had actually decreased to a negative result (suggesting no need for mental health referral). And in several cases the scores had dropped dramatically (e.g., from 80 to 64), but the second score was still above the cutoff (>62) and therefore the men's answers about score-change were technically correct. Another five men thought their scores had worsened, but in all these cases the score remained the same.

The vast majority of the men whose scores changed for the better were re-tested at the 5-month debriefing interview after a more substantial time in the batterer program or experiencing life changes. The ideal for a true reliability test would be within a couple of weeks of the initial intake

test, prior to the possible influence of the batterer program and additional treatments, and also beyond the potential impact of the court appearance and program intake.

## **BSI CONCURRENT VALIDITY**

### **Method**

In order to further test the concurrent validity of the BSI, the BSI scorings were compared to two additional screening instruments: one, the more comprehensive Psychiatric Diagnostic Screening Questionnaire (PDSQ; Zimmerman, 2002) that approximates DSM-IV-TR psychiatric diagnoses, and a brief instrument, called the Personality Assessment Screener (PAS; Morey, 1997), that focuses on personality traits associated with mental health problems warranting clinical attention (see Table 9). The PDSQ uses 111-items focused on psychiatric symptomology within the last 6 months, and the PAS is a 22-item instrument identifying problematic personality traits of the individual. The BSI contrasts with 53 items of distress within the past 7 days. The BSI symptoms are, therefore, more susceptible to being a situational reaction rather than a chronic condition. Also, the PDSQ includes 13 subscales (e.g. major depressive disorder, social phobia) and produces 6-levels of “dysfunction” or symptoms. The PAS assesses 10 different “elements” (e.g., negative affect, acting out, psychotic features) representing distinct domains of clinical problems and produces 5 levels of problems or symptoms. The BSI, on the other hand, is comprised of 9 subscales (e.g., depression, interpersonal sensitivity, anxiety) and a total global score that results in a positive-negative result, according to the test developers.

A subsample of 93 batterer program participants was administered the two screening instruments in addition to the BSI. A research assistant administered the battery of instruments to men at the end of batterer program sessions with groups of 8 to 15 men, until approximately 100 men

had completed the tests. This goal took about 6 weeks of different program groups to reach this goal. (Seven of the tests were discarded as invalid according to the scoring results.) The men completing the instruments had attended 7 to 10 sessions at the time of the screening. We thought that test administration at the sessions would help to avoid the apparent inflation of the BSI results at program intake revealed in our BSI re-test. Our initial analysis compared the screening results of the three instruments based on their recommended cutscores in two ways: 1) in terms of the portion of men who screened positive on each instrument, and 2) the extent of agreement across the three instruments. A second analysis assessed the association of the interval total scores for each instrument using Pearson's correlation. This analysis could reveal an association among the instruments that might otherwise be obscured by cutscore categories that were not necessarily equivalent.

## **Results**

The BSI results were closer to the PDSQ and PAS than we expected given the very high percentage of positively screened men with the BSI. According to the BSI cutscores, 47% of the men warranted further evaluation compared to 51% on the PAS (i.e., "marked" or "extreme" symptoms) and 55% on the PDSQ (i.e., average, high, and extremely high symptoms compared to psychiatric patients). There were, however, some discrepancies among the tests. Compared to the PDSQ, the BSI produced 8% false positives and 8% false negatives. Compared to the PAS, 6% of the BSI cases were false positives and 21% were false negatives. The PDSQ and PAS produced some discrepancies between themselves as well. The PAS had 16% false positives and 15% false negatives compared to the PDSQ. In sum, the BSI did not miss many cases that would warrant a further evaluation compared to the results of the PAS (8% false negatives), but may be missing as much as a

fifth of the cases warranting clinical attention according to the PDSQ ratings of moderate or high symptomology.

An examination of subscale agreement was limited to very few comparisons since most of the subscales or “elements” of the instruments did not directly match across all three instruments. However, 6 of the 9 BSI scales corresponded to subscales on the PDSQ. There appeared a fairly high agreement across the similar subscales overall. On the depression subscales, the BSI “depression” results agreed with 79% of the PAS results for “suicidal thinking,” and 81% of the PDSQ results for “major depressive disorder.” The BSI results for obsessive compulsive, somatization, phobic anxiety, and generalized anxiety agreed with approximately 80% of the PDSQ results for subscales of the same description. Both the PDSQ and PAS showed nearly half of the men having “alcohol problems” or “alcohol abuse,” similar to our previous results on the MAST at this same site (56%). With regard to extreme symptomology, only 4% of the men scored at the “extreme” level of the PAS, and 6% scored at the “high symptoms” or “extremely high symptoms” level on the PDSQ.

The correlation of the instruments’ total scores confirmed a substantial association among the three instruments. Pearson’s  $r$  ranged from .66 to .77 ( $p < .01$ ). The BSI total score (i.e., the GSI subscale) and the PAS total score were correlated at  $r = .67$  ( $p < .01$ ), the BSI was correlated with the PDSQ at  $r = .77$  ( $p < .01$ ), and the PDSQ and PAS were correlated at  $r = .66$  ( $p < .01$ ).

In sum, the BSI results substantially conformed to the results of complementary assessment instruments not only in terms of overall identification of “positive” subjects warranting referral, but also in terms of the results of similar subscales. There is however some disagreement among the instruments on “positive” subjects that ultimately should be further investigated. For instance, do the psychological profiles of those individuals (according to the various subscales) reveal a particular neglect or over-reach of the BSI that could be adjusted? Or is there a compilation of instrument



results that would optimize the screening—such as, using a positive screen on all three instruments to determine referral?

## **BSI PREDICTIVE VALIDITY**

### **Method**

The assumption of our research project was that batterer program participants with mental health problems were more likely to drop out of the batterer program and re-assault their female partners. As discussed in “Part I: Introduction,” previous research both in the domestic violence and alcohol treatment fields, appears to support this assumption. We might expect therefore that the BSI scorings would be associated with these outcomes—or “predict” them—if the BSI is measuring mental health problems. The predictive validity of the BSI might be tested against batterer program completion and victim-reported re-assaults during a 12-month follow-up period.

We examined the association of each of the nine sub-scales and global score index of the BSI to batterer program completion and victim-reported re-assault also to determine which type of mental health problems best predict program completion. First, the positive/negative results for the BSI subscales were cross-tabulated with program completion and re-assault from the 12-month follow-up, and logistic regressions were computed for program completion and victim re-assault using subscale scores as predictors after controlling for background characteristics. The BSI subscale positive/negative cutscores and also the BSI raw scores were entered using forward, backward, and direct enter procedures in order to confirm results. Additionally, variables for the number of positive BSI subscales and clustered BSI subscales (4 cluster solution with K-Cluster) were developed and entered into the equation as possible predictors. Lastly, “Receiver Operating Characteristics” (ROC)

were calculated for each BSI subscale to further test their prediction weighing the sensitivity and specificity of the scales in identifying program completion and re-assault.

As explained in “Part II: Research Methods,” completion of the required 4-months of weekly batterer program sessions was determined from the program attendance records for the full sample of positively screened men (N=479) and a systematic random sample of the negatively screened men (N=188). This allows for a greater range of BSI scores in the analysis of program completion. However, there is no re-assault information for any of the negatively screened men. The partners of the negative screens were not followed, since the initial research design focused on a comparison of the men under different referral options (i.e., voluntary and court-mandate). The predictive validity, in this later case, is not therefore of BSI positive/negative screen for re-assault, but of the BSI subscale scorings within the referred men.

### **BSI and Program Completion**

A cross-tabulation of BSI subscales with program completion shows that a positive score on the hostility and paranoid-ideation subscales is significantly associated with program dropout. The positive screens on these subscales were from 20-25% more likely to dropout than the negative screens. Specifically, nearly half (46%) of the positive screens on the hostility subscale completed the program compared to 63% of the negative screens completed ( $p<.01$ ;  $n=468$ ); and 53% of the positive screens on the paranoid-ideation subscale completed the program, compared to 63% of the negative screens ( $p<.05$ ;  $n=468$ ). There was no significant difference for the scorings on the other subscales. Although not statistically significant, those with several positive subscales (6 or more) were less likely to complete the batterer program ( $>5$  positive subscales=51% completion vs.  $<6$  positive subscales=59% completion.)

The logistic regressions controlling for batterer characteristics partially confirmed the cross-tabulations of the BSI subscales. The only significant BSI scales predicting program completion and dropout were the obsessive-compulsive, psychoticism, and paranoid-ideation subscale (at  $p < .05$  significance level). Interestingly, positive scoring on the obsessive-compulsive ( $OR = .334$ ) and psychoticism ( $OR = .344$ ) subscales decreased the odds of program completion, while positive scoring on paranoid-ideation increased the odds ( $OR = 4.71$ ). The hostility subscale may have dropped out of the equation because of its high association with violence-related variables entered as controls (e.g., prior domestic violence arrests, previous severe domestic violence, and how safe the women feel).

The significant control variables varied slightly across the logistic regressions for the outcome variables, but each equation included a variable related to relationship status and a variable associated with prior violence. Education also appeared in most of the regression equations as significant. As mentioned, ROC analysis with raw interval scores produces curves that might expose alternative cutscores. The ROC analyses for the BSI subscales, however, show a weak prediction for both the hostility ( $AUC = .564$ ) and paranoid-ideation ( $AUC = .565$ ) subscales, as well as the other subscales ( $AUC$  ranged between .483 for depression and .548 for phobic). There is no apparent advantage in an alternative cutscore.

### **BSI and Re-assault**

Only three of the BSI subscales were significantly associated with the victim-reported re-assaults during the 12-month follow-up (following batterer program intake). Men who scored positive on the BSI hostility subscale and paranoid-ideation subscale were more likely to re-assault their partners (Hostility: negative 29% vs. positive 40%;  $p < .05$ ;  $n = 293$ ; Paranoid: negative 24% vs. positive 37%;  $p < .05$ ;  $n = 293$ ). However, men were *less* likely to re-assault if they scored positive on

the BSI anxiety subscale (39% vs. 24%;  $p < .01$ ;  $n = 293$ ). (These two subscales were also significantly related to program dropout.) Notably, the GSI (Global Severity Index) and the number of positive subscales were not related to re-assault, or to program completion.

The logistic regressions controlling for background characteristics confirmed anxiety ( $OR = .43$ ;  $p < .01$ ) and paranoid ideation ( $OR = 1.83$ ;  $p < .05$ ) to be significant BSI predictors for cumulative re-assaults. As with program dropout, the hostility subscale was again not a significant predictor when controlling for the batterer characteristics that includes other indicators of hostility, specifically variables for violence and abuse. A positive score on the anxiety subscale reduced the likelihood of re-assault, whereas paranoid-ideation increased it. Not married, living with partner, and prior domestic violence arrest were also significantly related to the outcome.

The ROC analysis showed, however, the hostility subscale ( $AUC = .530$ ) and paranoid ideation subscale ( $AUC = .532$ ) to be the strongest predictor of re-assault among the 9 subscales, but still relatively weak in its prediction. The AUC for the other subscales ranged from .495 for depression to .419 for panic anxiety.

## **SUMMARY**

The BSI screening at batterer program intake identified an unusually high portion of men—over 40%—as warranting mental health evaluation and treatment. The characteristics and clinical diagnoses of the “referral sample” raised some concerns about the validity of these results. The referral sample was more likely to appear more at risk for program dropout and re-assault than those batterer program participants who scored negative on the BSI. The ultimate goal of the screening was to provide supplemental treatment to these men and improve these program outcomes. However, the clinical diagnoses of the small portion of men who did comply with the mandatory referral ( $n = 38$  of

148 or 26%) suggested that nearly half did not warrant mental health treatment. They had been diagnosed with adjustment disorders that could, however, been the result of the men not cooperating with the evaluation as well as a need for further clinical observation.

A further investigation of the reliability and validity of the BSI extended this contradiction. A test-re-test of the BSI (n=98) resulted in nearly half the men who initially scored positive scoring negative in the re-test, and the small group of men who initial scored negative (n=8) receiving a positive score on the re-test. A change in circumstances and the impact of the batterer program may have contributed to the change in test results. Most of those with different results took the re-test after a 4-5 month lapse. The initial test results may excessively over-screen program participants as positive due to the men's situational response to the stress of the court action and batterer program intake. The BSI is, after all, focused on current symptoms of distress.

Our examination of concurrent validity using the PDSQ and PAS, on the other hand, appears to endorse the screening results of the BSI (n=93). The percentage of men who screened positive on the PDSQ and PAS was similarly as high as the percentage with positive results on the BSI. And there was substantial agreement across the BSI screening and PDSQ and PAS screening results, as well as high agreement across similar subscales on the instruments. The BSI did, however, produce 8% false positives and 8% false negatives compared to the more comprehensive and diagnostic PDSQ. Some combination of instrument results, along with other clinical information, might be a more appropriate criterion for the referral.

The predictive validity of the BSI scores offers some additional considerations. The overall Global Severity Index (GSI) of the BSI was not associated with batterer program completion or re-assault during the 12-month follow-up; nor was the total number of elevated subscales associated with these outcomes. The GSI is however the primary indicator of a "positive" screening result. The

hostility and paranoid-ideation subscales, however, both reduced the likelihood of program completion and increased the likelihood of re-assault. The BSI anxiety subscale showed the opposite: more program completion and less re-assault. (The anxious men may be more wary of further sanctions and more sensitive to the program oversight.) These subscale results were not, however, entirely consistent when controlling for batterer characteristics and previous behavior. Indicators for “stake in conformity” were much stronger predictors of the outcomes. These results overall do suggest that a selective profile of BSI subscales, with an emphasis on hostility and paranoid-ideation, may be a more useful referral assessment. The question is whether men with such a profile would comply with an additional referral and treatment, especially after being more likely to dropout of the batterer program.

In sum, the BSI screening produces a high portion of men that do appear to warrant further treatment or intervention. Administering the screening at a later point, combining the screening results with other instruments, or using a profile of the BSI subscales might, however, reduce over-screening and excessive referrals.

## **BATTERER RESPONSE TO MENTAL HEALTH SCREENING**

### **INTRODUCTION**

We were also interested in the response of the batterer program participants to the mental health screening and referral. What were their perceptions of the protocol and screening? In their view, why did they comply or not comply with the referrals? What would they recommend to improve the process? To answer these questions, two interviews were conducted with the batterer program participants: 1) a case-management interview conducted with men under court-mandate for supplemental mental health treatment during the court mandated stage of the research project, and 2)

a debriefing interview at 5-months after program intake with all the men who screened positive on the BSI. The case-management interviews were conducted within 2-3 weeks of the program intake in order not only to collect additional information from the men, but also to promote compliance to the mental health referral. The debriefing interviews at 5-months after program intake were conducted when the men would have completed the required 16 sessions of the batterer program and a court review of their compliance to the batterer program and mental health referral.

A case-manager, under the direction of the principal investigator, interviewed 133 men out of 196 for a response rate of 67% as part of the case-management interviews. The response rate for the 5-month debriefing interviews was 53% (N=254 of 479). The completed debriefing interviews were with 76 men who were under voluntary referral, 90 during the transitional referral stage, and 88 men who were under mandatory referral for mental health evaluation and possible treatment. A comparison of demographic and behavioral characteristics assessed at program intake suggests a slight response bias in favor of men still in contact with their partners and less severely violent in the past.

The interviews asked about the men's compliance to the batterer program and mental health referral, reasons for not complying, expectations for mental health treatment, symptoms of mental health problems, and ratings and recommendations for the mental health evaluation and treatment. A tabulation of the main open-ended and closed-ended questions follows. Additional analyses compare the men's self-reports of compliance to clinical records in order to assess the validity of the men's reports and perceptions, and also compare individual's reports during the case-management interview to the later debriefing interview. Logistic regressions were computed in an attempt to identify predictors for referral compliance. Lastly, the following sets of variables were entered as predictors using a forward step procedure: demographics, prior behavior, other interventions, expected benefit

from treatment, and BSI subscales. Obtaining an evaluation and then receiving treatment, according to clinical records, were used as the outcome measure in two separate equations.

## **RESULTS AND DISCUSSION**

### **Referral Compliance**

**Results:** The compliance rates substantially varied across the men's interviews and clinical records. They increased over time from the earlier case-management interview to the later 5-month follow-up interview with the men; however, the clinical records showed a substantially lower compliance rate than the men reported in either interview (see Table 10). According to the clinical records, approximately 30% of the referred men obtained an evaluation and 20% received treatment during the mandatory referral. Moreover, there was a marked decline in the compliance to the steps of the referral (e.g., contacting the clinic for an appointment, attending the clinic for an evaluation, etc.).

The case-management interviews at 3-4 weeks after batterer program intake suggest a "half-life" of the referral compliance under the court mandate (see Figure 3). That is, about half (50%) of the men who screened positive made an appointment for an evaluation, half of those men (27% of mandated referrals) completed an evaluation, about half of those men (13% of mandated referrals) were recommended for treatment, and less than half of the recommended men actually received some mental health treatment (or 9% of the total mandated referrals). The men's 5-month follow-up show some increase in the compliance rates since the initial case-management interview, according to the men. In particular, the portion of men who reported obtaining an evaluation almost doubled from 27% to 55%, and those who eventually received some treatment increased from 9% to 24% (see Figure 4).



According to the *clinical records*, about half as many men made an appointment as said they did at the 5-month follow-up interview (clinical records 36% vs. man's report 58%) during the mandatory stage, and half as many actually obtained an evaluation (28% vs. 55%). About a fifth (19%) of the men received treatment according to the clinical records, compared to a quarter (24%) reporting so. Only 5% completed 10 sessions or more, according to the clinical records versus 17% in the 5-month follow-up. (These results remain constant with only those men who responded to the interviews and had clinical records available.)

**Discussion:** The discrepancy between compliance reported by men and their records raises the question: Why are the compliance rates derived from the clinical records substantially lower. The answer may lie in the reliability of the men's interview questions as well as some incomplete records. We did detect some confusion in what constituted "making an appointment" for an evaluation. When probed, at least a few men conceded that they had merely called the clinic and left a message. One man acknowledged that he thought the screening questions asked on his initial phone call to the clinic were an "evaluation." Moreover, some men could be exaggerating their compliance in an attempt to avoid any consequences for not complying. Our assumption is that the clinical records are the most accurate, but a few men may have obtained treatment from other clinics or private psychologists outside our clinic record search thus lowering the recorded compliance.

Although the portion of referred men receiving treatment remains low, there appears to be an encouraging increase in compliance following the case-manager's call at 3-4 weeks after program intake. It may take longer to contact the mental health clinic and obtain an appointment, even though the men were instructed to do so within the first week. Or, the call from the case-manager may have prompted and facilitated the additional compliance. Also, compliance increased from the transitional stage to the mandatory stage when the case-manager was implemented, according to the men's

reports in the 5-month follow-up. The percentage of men obtaining an evaluation went from over a third (37%) to over a half (55%).

As mentioned in “Part III: Formative Evaluation,” the greatest shortcoming may have been in the court support of the referral. Most all of the non-compliant men (97%) reported that the court did “nothing” in response to their non-compliance. Regardless of the low compliance, 88% of the men said that the batterer program should continue to screen and refer men to mental health services.

### **Men’s Expectations and Symptoms**

A substantial portion of the referred men expected some benefit from mental health treatment and identified symptoms to go with that expectation. At batterer program intake and the case-management interview, approximately one-third (34%-36%) of the positively screened men indicated that they would benefit from mental health treatment (see Table 11). A quarter (25%) continued to say that they needed help for mental health problems at the men’s 5-month follow-up. Over half (59%) at the case-management interview, and nearly three-fourths (72%) at the 5-month follow-up, identified symptoms that might have led to their referral. The most common symptom was “stress” (14% during case-management and 25% at the follow-up). Only a small portion (9%-11%) indicated alcohol problems. Interestingly, the men who did report symptoms at the 5-month follow-up reported a greater variety of symptoms than the men at the early case-management interview. (Perhaps the men were more aware of the extent of their problems after involvement with the batterer program and the mental health clinic.)

Those who identified benefiting from mental health treatment (34% of the batterer program intake sample) were, not surprisingly, 60% more likely to obtain a clinical evaluation (39% vs. 22%;  $p < .05$ ;  $n = 142$ ) and to receive mental health treatment (29% vs. 17%,  $n.s.$ ;  $n = 142$ ). In other words, the

self identification of “need” was a better predictor of compliance than the screening instrument results.

The vast majority (86%) of the men who did receive treatment (n=38) reported that it was helpful to some or a great extent, and over half (60%) said it helped them with their participation in and compliance with the batterer program. Two-thirds of the men (65%) interviewed at the 5-month follow-up admitted to needing other assistance in areas such as job placement (12%), couples counseling (12%), and housing (16%) among others. Probably what is most puzzling is that, despite the relatively low compliance, still 88% of the men at the 5-month follow-up agreed that the mental health referrals should be continued.

In sum, at least a third of the referred men perceived a need for their referral for mental health treatment and foresaw some benefit from it. This portion of men was, as one might expect, more likely to comply with the referral suggesting the role of personal incentive in the compliance. Most of the men who received treatment did view it as helpful and a substantial portion also wanted assistance in other areas. There does appear to be at least a small portion of men who need and *want* mental health treatment and other assistance—and who appear to benefit from it. The mandatory referral seemed to prompt more of these receptive men to comply compared to the especially low compliance under the voluntary referral. Personal incentive and external promptings may reinforce one another, but still to a relatively low result.

### **Compliance Predictors**

Our logistic regressions attempting to identify predictors of compliance revealed only one substantial predictor. Being on probation for a criminal offense increased the likelihood of *obtaining an evaluation* by 3½ times over not being on probation (OR= 3.56; p<.05). (Only one-quarter of the

men in our sample were on probation, since most of the men were ordered to batterer counseling under a stipulation of bond.) Not being married (OR=.39;  $p<.05$ ), having children living with you (OR=.36;  $p<.05$ ), having a protection order currently against you (OR=.33;  $p=.07$ ), and indicating little or no benefit expected from mental health treatment (OR=.49;  $p=.13$ ) reduced the likelihood of an evaluation by over 50%, but the last two of these variables only approached statistical significance ( $p<.15$ ) with our relatively small sample of mandatory referrals ( $n=148$ ).

Our efforts to identify predictors for *receiving treatment* produced different demographics with severe abuse substantially increasing the likelihood of treatment (OR=4.38;  $p<.05$ ). Older men were much more likely to receive treatment (>40 years old, OR=5.15;  $p<.01$ ), but African-American men (OR=.18;  $p<.01$ ) and men who see their partner only occasionally (less than daily, OR=.21;  $p<.01$ ) were less likely to obtain treatment. An expected benefit from treatment and a positive scoring on BSI subscales were not significantly related to treatment.

The strong association of “being on probation” with obtaining an evaluation may suggest reinforcement for compliance from the additional oversight from probation officers, and the “expected benefit from treatment” may suggest some personal incentive effecting compliance as well. The very different demographic predictors of treatment may reflect some men not cooperating at the evaluation and thus not being recommended for treatment. Particularly the younger men and African-American men appear to be more suspicious and resistant to the evaluation process. Interestingly, none of the BSI subscales, including hostility, were associated with obtaining an evaluation, nor was the number of elevated subscales. However, “prior incidents involving severe violence” was associated with receiving treatment.

## SUMMARY

Compliance with the steps of referral decreases with each step of referral (e.g., contacting the clinic, obtaining an evaluation, receiving treatment). Some of this decrease is explained as problems in contacting the agency to schedule an appointment or not receiving a diagnosis warranting treatment. The compliance appears to be related to several other factors, as well. One, compliance did increase following a call from the case-manager. Two, the men who expected a benefit from treatment were more likely to obtain an evaluation and treatment. Three, men under the additional supervision of probation were more likely to obtain an evaluation. Four, men who were likely to cooperate with an evaluation tended to receive treatment. There was no evidence that the BSI scorings were associated with compliance. There is, in sum, evidence that both personal incentive and external supervision contribute to compliance. The still low compliance may be further related to the men's report that the court did not further sanction them for non-compliance. Leverage from the supervision was missing. We can only speculate that leverage in the form of sanctions for non-compliance would have increased the rate of evaluations and treatment.

## ENDNOTES

1. One reviewer of this report interpreted our comparison of BSI scores to other scales and particularly the clinical diagnosis as an attempt to discredit the mental health problems evident among the batterer program participants. As discussed in this section, the high prevalence of men scoring positive on the BSI warranted further investigation for several reasons, rather than a personal bias or assumption. One, the BSI prevalence was almost double the results on the more sophisticated MCMI-III that more specifically approximates DSM-IV diagnoses (Gondolf, 1999b; White & Gondolf, 2000). Two, we were confronted by the very small portion of men receiving a clinical

diagnosis that warranted treatment. Three, the comparison to the PDSQ and PAS did show lower portions of men warranting referral based on 4-5 level categorization as opposed to the two-levels of the BSI.

2. One of the reviewers of this report argues that the analysis discussed in this section is not truly about “predictive validity” since comparing the BSI to the clinical diagnoses is based on the flawed assumption that the clinical diagnoses are more valid than the BSI. The BSI may, in fact, be more valid than the clinical diagnoses which are susceptible to a variety of confounding factors. The skill and experience of the clinician, the admission requirements of a clinic, the attitude of the client or patient, and the circumstances of the client’s appointment have all been shown to influence the diagnoses indicated by a clinician. However, the BSI is admittedly only a screening instrument of symptoms and not a diagnostic instrument as such. The concurrent validity of the BSI, as indicated in its comparison with other scales, is in question, as is its reliability over time. We termed our test “predictive reliability” with clinical diagnosis as our ultimate outcome, because the objective of the BSI screening is to identify men who are going to warrant mental health treatment. It is the clinical diagnosis that determines whether the men in our study reach the objective of receiving mental health treatment and is therefore the standard of the comparison in this intervention study.

3. A reviewer of the report also questions our identifying the BSI re-tests as a study of “reliability.” The reviewer points out that reliability more accurately refers to “ability of the instrument to consistently measure the attribute on different occasions,” not about scores remaining constant or stable over time. The issue of temporal stability, which is what our re-tests address, is but one element of the larger construct of reliability. Our stability study may, moreover, be confounded by the BSI’s narrow 7 day timeframe (i.e., the response is for symptoms in the previous 7 days); therefore, a different timeframe is assessed at each re-test. (As mentioned in endnote 2, the reviewer

also suggests that scales such as the BSI are likely to be more valid than clinical diagnosis and therefore more constant over time.) On the other hand, our concern was with the high percentage of positive screening, and whether the timing of the test influenced the results. The reviewer appears more focused on the construct or psychometrics of the instrument, or the stability of mental health problems assessed by the BSI. We wanted to know if circumstances of the testing contributed to the high rate of positive screening. Despite the qualifications that in terms of the long duration between the initial tests and some of the re-tests, we did find some evidence that the timing of the initial test may be a factor in the screening results.

## **PART V: BATTERER PROGRAM OUTCOMES**

### **INTRODUCTION**

The main objective of the research project was to test the impact of supplemental mental health treatment on batterer program outcomes. Our expectation was that the supplemental treatment for men warranting evaluation and treatment would improve outcomes in terms of increased batterer program completion and decreased re-assault of the men's female partners. As discussed in "Part II: Research Methods," the impact of mental health treatment was assessed for both "intention-to-treat" and "dose-response." The intention-to-treat was represented by the referral stages that ranged from voluntary referral to mandatory referral, along with a transitional stage. The dose-response was represented by the men who complied with the mandated referral by obtaining a mental health evaluation and receiving treatment.

The principal outcomes of batterer program completion and the men's re-assault were assessed as part of a one-year follow-up from batterer program intake. "Batterer program completion" was determined from the attendance records of the batterer program. These records were obtained for all 479 men who screened positive on the BSI and were referred for a mental health evaluation and possible treatment, and also for a systematic random sample of 188 men who screened negative on the BSI and were not referred. Attendance to the minimum requirement of 16 or more weekly batterer program sessions was considered "program completion." (This completion also represented successful adherence to the related program conditions outlined in Part II: Research Methods.)

"Victim Re-assault" was measured as a report of one of the physical aggression tactics on the Conflict Tactics Scales used against the men's female partners during any of the periodic interviews



—that is, the cumulative re-assault rate for the full one-year follow-up period. Other measures for non-physical abuse and women’s well-being were used as supplemental outcomes. The female partners of the positively screened and referred men were interviewed at program intake and every three months for a full year (i.e., program intake plus four follow-up interviews). A response rate of 65% was achieved for the full 12 months. A detailed description of the follow-up interviews and procedures appears in “Part II: Research Methods.”

The analysis consists of a cross-tabulation first of the outcome measures by the intention-to-treat indicator and then the dose-response indicator, and the rates across the treatment options were compared using Chi-Square or Fisher Exact Test ( $p < .05$ ). As also explained in “Part II: Research Methods,” a multivariate confirmatory analysis was conducted using logistic regressions for both outcomes. In order to control for subsample characteristics that may influence the cross-tabulation results, batterer demographics, relationship status, prior behavior, previous interventions, and treatment perceptions were entered stepwise as blocks into the regression followed by the direct entry of either the indicator for intention-to-treat (i.e., the referral stages) or for dose-response (i.e., obtaining an evaluation and receiving treatment). If either of the treatment indicators significantly predicted one of the outcomes, we would assume a “treatment effect” beyond the possible influence of batterer characteristics.

Re-arrest rates for a subsample of subjects were used to further confirm the findings of the re-assault analyses for intention-to-treat (i.e., voluntary vs. mandated referrals). Additionally, “re-arrest” offered an outcome measure, besides program completion, for the batterer program participants who were *not* referred to a mental health clinic, and thus enable a further comparison of the referred men to those men not referred. Arrest records were sought for a random subsample of 300 of the original 1043 program participants screened at program intake, and the obtained records

were coded for the types of crime involved in the arrests (i.e., domestic violence, other violence, alcohol and drug offenses, and other crimes). Details of the subsample, coding, and analysis are discussed, along with the results, in the section on “re-arrests” below.

## **RESULTS FOR PROGRAM COMPLETION**

Our first expectation in terms of outcome was that mandated referral and compliance would improve batterer program completion rates for the men who screened positive for mental health problems. The overall completion rate for the entire sample of positive screens (N=479) was 57% for the required 4 month weekly program. This completion rate was slightly lower than the 62% rate for a systematic random sample of the men who screened negative on the BSI screening (N=188). The same rate of both positive and negative screened men completed at least 3 sessions or more of the program (84%). The completion rate for a 1995 sample at this same site was 69%, but the program at that time was only 3 months long.

We first considered the intention-to-treat effect on program completion as represented in the referral stages (i.e., voluntary, transitional, and mandatory) (see Table 12). The men under the mandated referral were slightly more likely to complete at least the first three sessions than men in the voluntary and transitional stages (voluntary: 80% and transitional: 83% vs. mandated 88%; n.s.; N=479). However, the program completion rate decreased slightly under the mandated referral compared to the transitional and voluntary referral, although not significantly (voluntary: 57% and transitional: 60% vs. mandated 52%; n.s.; N=479). This result may have been influenced by the stricter enforcement of participant payment.

The dose-response condition showed a statistically significant increase in batterer program completion for those men who obtained a mental health evaluation and those who received mental

health treatment (see Table 13). Specifically 76% of the evaluated men compared to only 38% of the non-evaluated men completed the batterers program ( $p<.05$ ;  $n=148$ ), and 68% of the treated men completed the program versus 45% of the non-treated during the mandated referral stage ( $p<.05$ ;  $n=148$ ). The relationship between the referral compliance and the batterer program completion is unclear, however. Many of the men who were intent on completing the batterer program may have felt compelled to seek a mental health evaluation, rather than the evaluation or treatment contribute to their completing the batterer program. Nonetheless, 60% ( $n=21$ ) of those who did receive treatment ( $n=35$ ) indicated that the mental health treatment was helpful to “some extent” or a “great extent.”

As outlined in “Part II: Research Methods,” logistic regressions were computed to control for possible differences in the characteristics of the referral stages and referral compliance subsamples that might influence the cross-tabulation for batterer program completion. The regression results appeared to confirm the cross-tabulations. In our test for intention-to-treat, the “referral stages” variables were not significant when directly entered into the regression following the stepwise entry of batterer characteristics (i.e., demographics, relationship status, prior behavior, prior interventions, and attitudes toward treatment). However, men who obtained a mental health evaluation were six times more likely to complete the batterer program, even after controlling for batterer characteristics ( $OR=6.30$ ;  $p<.001$ ), and men who received treatment were two times more likely to complete the program but only approaching statistical significance with the small number of men complying to treatment ( $OR=2.18$ ;  $p<.1$ ). In the regression including referral stages, men over 40 were more likely to complete the batterer program ( $OR=1.86$ ;  $p<.05$ ); also, men with some college education ( $OR=.47-.46$ ;  $p<.001$ ) and men with partners who were also arrested for domestic violence ( $OR=.50$ ;  $p<.05$ ) were less likely to complete the program. In the regressions for mental health evaluation and treatment, men with college education ( $OR=.43$ ;  $p<.001$ ) and with an arrested partner ( $OR=.43-.42$ ;

$p < .01$ ) were again less likely to complete the program. Men who were employed were also more likely to complete ( $OR = 1.68 - 1.64$ ;  $p < .01$ ), and men with a partner using drugs were less likely to do so ( $OR = .55 - .57$ ;  $p < .05$ ).

## **RE-ASSAULT AND ABUSE OUTCOMES**

### **OVERALL RE-ASSAULT AND ABUSE**

The second and most important outcome was the re-assault committed by the batterer program participants. We first examined the overall re-assault rates for our sample of men who screened positive for mental health problems. This included all the positive screens whether or not they completed the batterer program or complied with mental health referral (see Table 12). The re-assault rates reported by the men's female partners indicate a third (32%) of the men re-assaulted during the one year follow-up (0-12 months after program intake;  $n = 308$ ). Approximately a fifth (18%) of the men assaulted their partner during the first 3 months after program intake, and 9% re-assaulted during the last three months of the follow-up (9-12 months). Approximately, half of the men who re-assaulted at the last three-month interval had previously re-assaulted their partner in previous intervals of the follow-up. The re-assault rate for the 9-month period *following* the required number of batterer program sessions (3-12 months) was 25%.

As mentioned in "Part II: Research Methods," we considered other indicators for abuse, as well. The overall rate of threats during the 12-month follow-up was 36%; 14% of the men's partners reported "severe" assaults during the follow-up, and 18% reported being bruised or injured ( $n = 308$ ). We also considered the women's ratings of their well-being at 12 months. Three-quarters (75%) of the men's initial partners rated their life overall as "good" or "very good" at the 12-month follow-up interview ( $n = 292$ ). About two-thirds (62%) felt that it was "very unlikely" that their partners would

hit them again, and approximately the same percentage felt “very safe” (66%). The women’s perceptions reflect the reverse of the re-assault rates; that is, about 80% of the women who felt “very safe” had not been assaulted during the follow-up.

The research design did not include a follow-up of the victims with men who did not screen positive. However, we can compare the re-assault rates with rates at the 12-month follow-up with a sample of program participants enrolled at the same site in 1995 (Gondolf, 1999a). The tracking and interviewing methodology were nearly identical in the current and the previous studies. The overall rates and the trend of re-assault over the intervals is equivalent for the current sample of positive screens and the previous sample of all those referred to the program. The demographic and behavioral characteristics of the 1995 and current sample do not significantly differ except for the current sample reporting higher rates of unemployment, drug use, and alcohol use.

## **RE-ASSAULT AND ABUSE FOR INTENTION-TO-TREAT**

We expected that the intention-to-treat, represented by mandatory referral to supplemental mental health treatment, would result in lower re-assault rates for men who screened positive for mental health problems. As with batterer program completion discussed above, we compared the re-assault outcomes for the stages of referral (i.e., voluntary, transitional, and mandated) (see Table 12). There is a substantially lower, but not statistically significant, re-assault rate for the mandated referral group compared to the transitional referral group at the cumulative 9-month follow-up (mandated: 23% vs. voluntary: 34% and transitional: 29%; n.s.; n=329), but a similar rate across the referral stages for the full 12-month follow-up (mandated: 33% vs. voluntary: 32% and transitional: 30%; n.s.; n=308). Moreover, the re-assault rates following the batterer program (3-12 months) were also the same across the referral stages (25-26%).

Several other indicators for physical and non-physical abuse also show similar rates across the referral stages and no apparent effect for intention-to-treat under the mandated referral stage (see Table 12). This includes the cumulative 12-month follow-up rates for threats (mandated: 43% vs. voluntary: 36% and transitional: 43%; n.s.; n=307), severe physical abuse (mandated: 14% vs. voluntary: 16% and transitional: 11%; n.s.; n=307), and bruise or other injury (mandated: 18% vs. voluntary: 18% and transitional: 19%; n.s.; n=307). The woman's overall rating of their well-being at the 12-month follow-up was also similar across the three referral stages (mandated: 70% vs. voluntary: 76% and transitional: 75%; n.s.; n=292). However, the women reported a significantly lower sense of safety during the mandated stage both in terms of the likelihood to be hit ("very unlikely" to be hit" for mandated: 50% vs. voluntary: 67% and transitional: 66%;  $p < .05$ ; n=256 for the initial female partner at 9-12 month follow-up interval), and their feelings of safety ("feel 'very safe'" for mandated: 54% vs. voluntary: 73% and transitional: 69%;  $p < .05$ ; n=256). This decrease for the mandated stage remains constant when controlling for the woman's contact with the man (i.e., "see everyday" and "no contact"). These differences exist even though the mandatory women have less partner contact (no partner contact at 12 mos. for mandated: 12% vs. voluntary: 29% and transitional: 23%;  $p < .05$ ; n=308), and counter our expectation that the intention-to-treat would improve outcomes of the batterer program. It appears that the men were more hostile or threatening toward the women in reaction to the additional requirement of mandated referral.

## **RE-ASSAULT AND ABUSE FOR DOSE-RESPONSE**

The second question is whether compliance to the mental health referral, in terms of obtaining an evaluation and treatment, reduces re-assault (see Table 13). Those who completed an evaluation under the mandated referrals were nearly a third less likely to re-assault their partner during the full

12-month follow-up (evaluation: 22% vs. no-evaluation: 31%; n.s.; n=119), and a third less likely to re-assault if they received treatment (treatment: 19% vs. no-treatment: 30%; n.s.; n=119). (The small sample size contributes to this substantial difference not being statistically significant.)

Other indicators of abuse were also lower for those men obtaining an evaluation. For example, the evaluated men were approximately half as likely to have threatened their partners (evaluation: 27% vs. no-evaluation: 50%; n.s.; n=92) and to have “severely” assaulted them (evaluation: 9% vs. no-evaluation: 18%; n.s.; n=92), but nearly equivalent for the treated versus untreated men.

The differences in the women’s perceptions for the dose-response were not significant or offsetting (i.e., for men receiving a mental health evaluation or treatment). If the man obtained an evaluation, his partner was more likely to report that being hit was “very unlikely” and feeling “very safe.” However, a lower percentage reported further hitting was “very unlikely” and feeling “very safe,” if the man received treatment (see Table 13). The differences of 10-12 percentage points must be viewed with caution because of the small sample size (n=90) of follow-up responses during the mandated stage, and the consequential lack of statistical significance.

The dose-response outcomes, of course, could be influenced by the additive effect of batterer program completion. We therefore controlled batterer program completion in a three-way cross-tabulation, but found no significant effect of the program completion on the dose-response. The men who completed a mental health evaluation were approximately a third less likely to re-assault their partners both when not completing (no-evaluation: 31%/n=16 vs. evaluation: 17%/n=1; n.s.; n=58) and completing the batterers program (no-evaluation: 31%/n=10 vs. evaluation: 23%/n=6; n.s.; n=59). This tendency is less favorable for those who actually obtained mental health treatment. The treated men who completed the batterer program were 20% less likely to re-assault (5 percentage

points less than the no-treatment group; no-treatment: 28%/n=12 vs. treatment: 24%/n=4; n.s.; n=59), but those treated men who did not complete the batterer program were still a third less likely to re-assault (no-treatment: 33%/n=16 vs. treatment: 11%/n=1; n.s.; n=58). These results, however, are based on a very small sample (n=117) and are not statistically significant according to conventional significance tests (i.e., Fisher's exact test and Chi-square test).

## **CONTROLLING FOR BATTERER CHARACTERISTICS**

In the confirmatory logistic regressions, neither the intention-to-treat nor the dose-response variables were statistically significant predictors for re-assault during the 12-month follow-up. The odds ratios for the dose-response variables did, however, show at least a tendency in the expected direction of reducing the likelihood of re-assault by 50%, similar to the results of the cross-tabulations with re-assault by evaluation and treatment. The batterer characteristics significantly associated with re-assault were not being married (OR=2.31;  $p<.01$ ), living with a partner (OR=2.05;  $p<.01$ ), and having been previously arrested for domestic violence (OR=2.44;  $p<.01$ ). These predictors were consistent in both the intention-to-treat and the dose-response regressions. In sum, the regressions substantiate the results of the cross-tabulations showing no significant intention-to-treat effect and no dose-response effect on the batterer re-assaults. This might be expected given the very similar characteristics across the subsamples representing the intention-to-treat condition, but was less certain for the dose-response subsamples where the characteristics of treated and non-treated did differ (as suggested by the predictors of referral compliance discussed in a previous section).



## CONFIRMATION USING RE-ARRESTSS

### METHOD

We also examined the re-arrests of the subjects to further test for an effect from mental health referral compared to non-referrals, and also the effect of mandated referral compared to voluntary referral in our “intention-to-treat” approach. The arrest records, like the program attendance records, offered outcome information of the non-referrals that were not included in the follow-up interviews assessing re-assault. Re-arrests for domestic violence and other crimes have been frequently used in batterer intervention evaluations as an additional outcome, because they are considered by the court to be a measure of recidivism and therefore “failure” of a particular case. However, they tend to reveal fairly low base rates for domestic violence arrests (e.g., 8-10%) during a 12-month follow-up and are not necessarily representative of the actual level of violence and abuse. They are subject to variations in police practices and women’s confidence in calling the police. Therefore, while the arrest records offer a convenient outcome indicator and one of interest to the criminal justice system, their results must be viewed with caution and are used here as a confirmatory supplement to the re-assault results.

“Re-arrests” were determined by reviewing the arrest records provided by the state-wide database of criminal histories. The records include all arrests in the state even if the charges were dropped. Records were requested for a stratified random subsample of 300 men from our initial intake sample of 1043 men. This subsample included 100 men who were voluntarily referred for mental health evaluation and treatment, 100 men whose referral was mandated by the court, and 100 who were *not* referred because of negative results on the BSI and ADS. Records were available for 254 of the 300 subjects (15% were missing). The men with no records may have had their record expunged or used aliases that prevented locating their record. The characteristics of the men with

missing records were not significantly different than those of the men with records, suggesting that the missing records were not likely to bias the results.

The available records were coded for offenses involving domestic violence, other violence (i.e., violence toward non-family members or acquaintances), drug and alcohol-related crimes (including drunk driving and drug possession,) and other crimes (i.e., any crimes not in the previous categories) that occurred during the 12-month follow-up period. If the charges associated with a particular arrest included more than one category of crime, the arrest was categorized as the highest priority crime indicated by the above order of crime categories. These codes were tabulated across the 12-month follow-up period to also identify the occurrence of “any arrest” and the “total number of arrests.” The crime categories of re-arrest were then cross-tabulated, first, by the non-referred men (n=77) versus referred men (n=177), and, second, by the voluntarily referred men (n=86) and mandated men (n=91).

## **RESULTS**

A comparison of the re-arrest rates for the men referred to a mental health clinic versus those *not* referred shows no statistically significant differences between the two groups ( $p < .05$ ; see Table 14). However, the subsample of referred men were 40% less likely (8% vs. 14%;  $p = .17$ ;  $n = 254$ ) to be re-arrested for other violence (i.e., other than domestic violence). This difference appears even though the referred sample had a slightly higher portion of men who had been previously arrested for domestic violence and other crimes (domestic violence: 55% vs. 48%, *n.s.*,  $n = 254$ ).

A comparison of those men who were referred voluntarily versus those referred as part of a court mandate were similarly 40% less likely to be re-arrested for both domestic violence and other violence, but the difference was not statistically significant (domestic violence: 4% vs. 7%; *n.s.*,  $n = 177$ ; other violence: 6% vs. 11%; *n.s.*,  $n = 177$ ) (see Table 15). The mandated referrals were also

significantly less likely to be re-arrested for crimes other than violence or alcohol and drug offenses (other crimes: 3% vs. 12%;  $p < .05$ ;  $n = 177$ ), and as a result less likely to have been re-arrested for any crime during the 12-month follow-up (18% vs. 29%;  $p < .1$ ;  $n = 177$ ).

These comparisons of re-arrest rates appear to reinforce the results of other outcome measures. Similar to the comparison of re-arrests, we found little difference in program completion rates for the men referred to a mental health clinic and those who were not referred. The men identified for mental health treatment do not stand out as doing any worse than their non-referred counterparts, even though the rationale for identifying them is that they do tend to have less positive outcomes in batterer programs. This is especially the case for men who received no supplemental mental health treatment (i.e., voluntary referral). Interestingly, those men who were mandated to a mental health clinic for evaluation and possible treatment were actually less likely to be rearrested for violent or other crimes—a result counter to the re-assault rates for the 12-month follow-up.

The *re-assault* rates across the referral stages (i.e., voluntary, transitional, and mandated) suggested no effect from the intention-to-treat. The re-arrest rates must be regarded with caution given the low rates and insignificant differences, but they do at least suggest that the additional case requirements and oversight, and the increased contact with mental health clinics may have contributed to an improved outcome. Because of the additional oversight, the mandated men may have been more cautious about getting arrested for their offenses.

## SUMMARY

As discussed at the beginning of this chapter, the men referred to mental health treatment (i.e., those who screened positive on the BSI or ADS) appeared more predisposed to poor batterer program outcomes than their non-referred counterparts. They had lower “stake in conformity” (e.g., more

unemployment, less with children), heavier drug and alcohol use, higher violent behavior and previous arrests, and more prior mental health treatment. However, the referred men had similar rates of program dropout and re-arrests, rather than the higher rates expected for men with additional problems. This similarity may be due to over screening with the BSI that included many “normal” men in the referred group who helped to counteract the potential for higher re-arrest rates. More likely, the predisposition for higher dropout and re-arrest was suppressed or contained by the additional supervision and requirements, despite the vast majority of referrals not actually obtaining mental health treatment.

We did find some tentative evidence that supplemental mental health treatment might improve batterer program outcomes. The intention-to-treat, represented by the mandatory referral, did not significantly improve batterer program outcomes in terms of dropout, re-abuse, and re-assault. The mandated referrals did, however, have substantially lower re-arrest rates for violent crimes and especially for other types of crimes. Re-arrest rates as an outcome measure must be viewed with some caution since re-arrests may be influenced by perpetrators’ avoidance of arrests, victims’ reluctance to call police, and police response to calls. Therefore, the outcomes for intention-to-treat overall point to “no effect” for the mandatory referral and the treatment intention it represents.

The dose-response showed a substantial tendency in the expected direction. That is, men who obtained a mental health evaluation and men who received treatment were more likely to complete the batterer program and less likely to re-assault their female partners. This latter tendency must be considered tentative because it is not statistically significant and based on a very small number of compliant men. It is, furthermore, likely to be confounded by motivational differences and other characteristics that were not controlled in the comparison between those men treated and not treated. Even if the apparent treatment effect were replicated in a controlled comparison, the effect size would

be very small overall. In other words, any improvement in outcome would benefit only a small portion of the men entering a batterer program despite a large amount of referral screening and oversight (approximately 12 out of 200 program enrollees). The question remains, however, whether greater compliance would have an increased effect.

## **PART VI: SUMMARY AND DISCUSSION**

### **INTRODUCTION**

This research project investigated the implementation and effectiveness of screening and referring batterer program participants for supplemental mental health treatment. The aim was twofold: one, to examine the feasibility, appropriateness, and utility of screening with a widely used instrument, the Brief Symptom Inventory (BSI); and two, to test the impact of the referral on batterer program outcomes, especially completion of a four-month batterer program and re-assault during a one year follow-up. A series of studies were conducted to address these aims, including a descriptive formative evaluation of the implementation of mandatory referral, reliability and validity tests of the screening instrument and procedures, an examination of referral compliance and clinical diagnoses, and an outcome evaluation of mental health treatment both in terms of intention-to-treat and dose-response. Each section below summarizes findings from a separate set of studies and discusses implications for practice. The limitations of the research project and directions for future research appear after these summaries.

### **IMPLEMENTATION**

#### **SUMMARY**

One major over-riding finding was the difficulty in fully implementing the mandatory referral across the collaborating agencies—the courts, mental health clinics, and batterer program. Under “mandatory referral,” compliance to the mental health referral was considered part of the batterer program requirement and non-compliance was subject to further court sanctions. Despite extensive planning, training, protocols, and support; inconsistencies in the screening and referral emerged and

undercut the men's compliance. The addition of an experienced case-manager and system-coordinator substantially improved implementation and with it, referral compliance. This person worked directly under the supervision of the principal investigator and was therefore relatively independent of the individual agencies. She was able to monitor the "big picture" through the ongoing data collection of the research project, and remain relatively free of the demands and pressures within the agencies. Notably, the court response to the men's non-compliance continued to be inconsistent according to our observations and the men's reports. The leverage of the mandatory referrals was therefore diminished and a fuller compliance hampered. The main reasons for the court inconsistencies were reportedly the expansion of judges circulating through the domestic violence court and the tendency to move cases through the system to avoid a backlog.

The other major issue is the different missions of each of the agencies. They don't necessarily clash, but do pose different priorities, orientations, and emphases. One example lies in a third of the referred men being diagnosed with adjustment disorders and not being recommended for treatment. As staff from the mental health clinic explained, the clinicians are accustomed to voluntary clients and individuals who present psychiatric symptoms. Many of our referred men tended to introduce their relationship problems and to refuse to elaborate on their mental health issues. The practical demands in the clinic, as well, preclude special accommodations and outreach to the batterer referrals. The intake unit of the clinic, for instance, received up to 100 calls a day, and each clinician saw as many as 15 clients in various states of need.

## **IMPLICATIONS**

The difficulties with mandatory implementation raise further questions about so-called "coordinated community response" and the efforts to develop agency collaboration for domestic

violence cases. A series of studies suggest that increased coordination and collaboration tend to increase service delivery, referral compliance, and positive outcomes for batterer programs (e.g., Bennett et al., 2007; Gamache, Edleson, & Schock, 1988; Gondolf, 2000b; Murphy, Musser, & Maton, 1998). However, another set of recent studies exposes the shortfalls of coordination efforts and questions their effectiveness (California State Auditor, 2006; Labriola et al., 2007; Visher, Newmark, & Harrell, 2006). These studies include a national survey of agencies involved in batterer intervention; a state review of agency compliance to batterer intervention guidelines, and a demonstration project of enhanced coordination in domestic violence cases. They note many of the same difficulties encountered in our research project: training decay, staff turnover, diverted funds, competing priorities, and system overload.

The interpretation of these shortfalls is unclear however. Is “community coordinated response” an illusive ideal undercut by intractable “real world” circumstances? Or, has it not been fully realized due to a lack of adequate resources, organizational restructuring, and commanding leadership? The improved implementation under our case-manager and system-coordinator, amidst several unexpected challenges, suggests that coordination and collaboration can be improved with a modest investment. While the referral compliance and ultimately service delivery increased, it still fell far short of the goal. Moreover, the overall effectiveness of increased “services” in reducing domestic violence—i.e., supplemental mental health treatment—was negligible in our research project, as was the case in the demonstration project of enhanced coordination in domestic violence cases (Visher et al., 2006). It is, of course, very difficult to tease out the impacts of coordination community responses, and ultimately to determine the outcome of its full implementation.



## SCREENING

### SUMMARY

The mental health screening at batterer program intake identified nearly half the men as warranting a referral to a mental health clinic for an evaluation and possible treatment (N=479). This determination was made using the prescribed cutscores of the Brief Symptom Inventory (BSI); the Alcohol Dependency Scale, and referrals directly from the court added few additional cases. A third of the men who screened positive on the BSI scored the maximum on at least one of the BSI's subscales, and a third had scored positive on several subscales (>5 subscales). The men who scored positive, moreover, appeared at risk for re-assaulting their partners, according to their demographic and behavioral characteristics, and therefore warranted more extensive intervention or attention.

These screening results were surprisingly high, especially given that less than 25% of the men showed evidence of a major psychiatric disorder using the Millon Clinical Multiaxial Inventory (MCMI-III; Millon, 1994) in a previous study at the same site (Gondolf, 1999b). It also seemed high given the incongruence between the BSI results and the clinical diagnoses of the 38 men of 148 referred under a mandatory referral stage of the project. Over a third of these men were diagnosed with adjustment disorders and were not recommended for further treatment. Our initial suspicion was that the screening over-identified men as "positive" because of the instruments focus on distress within the past seven days rather than on chronic symptoms associated with mental health disorders. The BSI results may have reflected the men's reaction to the "distressing" circumstances of arrest, court action, and batterer program intake. We conducted a series of reliability and validity tests to confirm our screening and explore for alternative scorings or approaches.

A reliability re-test (n=98) showed that the majority of men who initially scored positive did score negative on the later test. Most of these men, however, were re-tested at 3-5 months after the

initial testing at program intake, and attributed their negative BSI scoring to a change in circumstances on the benefit from the batterer program. In other words, the distress symptoms registered on the initial BSI scores may have been reduced over time, rather than the initial results not having been reliable.

Our test of concurrent validity appeared to confirm the screening results of the BSI as well. A more extensive instrument linked to the DSM-IV (PDSQ) and a shorter instrument focusing on more ingrained personality traits (PAS) identified a very similar portion of men warranting mental health referral. The overlap among the instruments was, moreover, very high especially for the BSI and PDSQ. Similar subscales across the instruments also were highly correlated. Our test of predictive validity points to the importance of two BSI subscales. The BSI subscales for hostility and paranoid-ideation were associated with program dropout and also with re-assault during the one-year follow-up. The raw scores of these subscales were the strongest predictors of dropout and re-assault in ROC analyses, but relatively weak ones with no clear cutscore alternatives. The hostility subscale was not a significant predictor when controlling for batterer characteristics in a multivariate analysis of dropout and re-assault. Controlling for batterer characteristics of severe abuse, criminality, and drug and alcohol use are likely to have countered the influence of hostility in the analysis. Overall, it appears that the accusations and suspicions associated with hostility and paranoid-ideation warrant special attention and may be reflected in the resistance to clinical evaluations and the prevalence of adjustment disorder diagnoses.

## **IMPLICATIONS**

The reliability and validity tests offer a degree of confirmation for the BSI screening results. There is some tentative evidence that screening subsequent to the batterer program intake might

reduce over-screening, along with combining the screening results with other instruments or using a profile of the BSI subscales. Some consequences accompany these alternatives however. A delay in screening is likely to miss many men who warrant the additional help at the outset and might otherwise dropout before such a screening. Administering multiple instruments can also be perceived as a burden to the men and raise their resistance to the testing. Multiple instruments also take more time and supervision, especially with men who generally have low attention spans and educational deficits. There is no doubt that a profile that includes hostility and suspicion are associated with poor outcomes, but it is less clear what levels of hostility and what combination of problems warrants supplemental treatment. Also, hostile men are generally not good candidates for treatment; they tend to be non-compliant and uncooperative.

One other alternative is, of course, to do a full clinical evaluation of each individual man entering the batterer program, as some batterer programs already do. Such an evaluation might also help to engage a man, as well as offer a fuller picture of his needs. The main consequence here is the substantial costs in time, money, and staffing to conduct these sorts of evaluation. The vast majority of batterer programs are currently under-funded and struggling to sustain their current level of service. There are nearly 2000 men a year referred to batterer programs in the Pittsburgh area of a half-million population. Individual evaluations would amount to an additional 2000 hours of a clinician's time at a total cost of about \$80,000 to \$100,000 in salary and facility. At the same time, our current research and several previous studies raise question about this extensive investment. It is still somewhat unclear exactly what to identify as detrimental to program outcomes, and what impacts supplemental or alternative treatment might have.

## **REFERRAL**

### **SUMMARY**

The referral compliance was overall relatively low, especially given the agency procedures and coordination involved in the research project. Under the mandatory referral stage, approximately a fifth of the referred men obtained a mental health evaluation and received treatment, and a portion of these men were already currently in some form of additional treatment. That is, less than 10% of the men entering the batterer program received supplemental mental health treatment. According to follow-up interviews with the men, most of the treated men reported having mental health problems and benefiting from the treatment both in terms of mental health and domestic violence. Despite the relatively low compliance, still 88% of the men at the 5-month follow-up recommended that the mental health referrals should be continued.

Compliance appears, not surprisingly, to be related to external reinforcements and personal incentive. Besides the self-identification of treatment need, the additional oversight from probation officers and a man's expecting some benefit from treatment were associated with compliance. The introduction of mandatory referral with sanctions for non-compliance and of a case-manager with promptings and encouragement also substantially increased compliance. Interestingly, the BSI subscale scorings did not predict compliance. That is, the men who especially warranted treatment, according to the screening, did not necessarily receive it. The implication might be that an externally reinforced referral for the men who self-identify a need and interest for supplemental treatment may be more efficient than screening with the BSI and the complicated protocols it requires.

The implementation of the mandatory referral illustrates the influence on compliance of what has come to be called "accountability"—that is, consistent oversight, supervision, and coordination of required actions and sanctions for not meeting those requirements. In our own previous studies, "no-

shows” of court referrals to the batterer program dropped from 30% to 5% (with a consistent completion rate of 70%) when additional court review and sanctions were introduced (Gondolf, 2000b). In the current research project, only about 5% obtained a mental health evaluation under voluntary referral, compared to 28% under mandatory referral that included a case-manager, according to clinical records. A case-manager and system-coordinator, moreover, enhanced the implementation of a transitional stage prior to the more fully implemented mandatory referral. The men’s compliance also increased following the case-manager contact, according to the men’s follow-up interviews. (This increase may be in response to the case-manager, or to delays in getting appointments at the mental health clinic.)

One additional observation is the progressive attrition in compliance across the referral steps. According to case-management interviews at 3-4 weeks, about half (50%) of the men who screened positive made an appointment for an evaluation, half of those men completed an evaluation, about half of those men were recommended for treatment, and less than half of the recommended men actually received some mental health treatment. Only half of the treated men attended the recommended minimum of 10 sessions. The complexity of the referral compliance, in terms of so many steps to mental health treatment and the additional requirements of batterer program participation, appears as a barrier in itself. There is at least some research suggesting that more simplified procedures would improve compliance. We previously found, for instance, that batterer program referral from a preliminary hearing and court supervision of batterer program compliance had advantages over more complicated criminal court referral and probation supervision (Gondolf, 1999c).

## **IMPLICATIONS**

These findings point us to an alternative approach for supplemental treatment—one that would reduce referral steps and also addresses agency inconsistency and lack of integration. At least two batterer programs, one in Calgary, Canada, and the other in Fond du Lac, Wisconsin, are experimenting with hospital-based domestic violence programs. Men identified with mental health problems participate in a program that includes mental health treatment and domestic violence education together. Additional individual counseling sessions may be included based on clinical evaluations or observation. Another approach used in some batterer programs, such as AMEND in Denver, is to provide individual treatment for mental health problems at the batterer program agency to complement the group sessions on domestic violence. Both approaches consolidate referral procedures and also integrate the mental health treatment. The major criticism is that especially the hospital-based approaches may dilute the domestic violence education and also inadvertently redefine domestic violence as a psychological problem. These approaches also add another layer of services and the accompanying costs, rather than take advantage of the existing services and agencies as cross-agency referral does.

## **OUTCOMES**

### **SUMMARY**

The crucial question of the research project was whether the supplemental mental health treatment would improve batterer program outcomes. The answer is somewhat mixed, but not encouraging overall. The intention-to-treat under the mandated referral showed little improvement over voluntary referral and a transitional referral stage of inconsistent oversight and no case-management. There was no significant difference in both batterer program completion and re-assault

of the men's female partner during the 12-month follow-up. This finding persists when controlling for batterer characteristics.

There were also no differences for other forms of abuse and the women's sense of well-being, except that the women's sense of safety and likelihood of being hit was significantly *less* among the women during the mandatory phase. It could be that some men become more resentful and hostile in reaction to the referral requirement, and that hostility is observed and felt by the women. However, re-arrest rates for violent and other types of crimes were substantially less for the mandatory referrals compared to the voluntary referrals. The additional supervision and accountability under the mandate may have helped to suppress reportable crimes or women's response to them.

The dose-response offered some tentative evidence that complying with the referral and actually receiving treatment improved outcomes. This suggestion must be viewed with caution for a number of reasons. Over two thirds of the treated men completed the batterer program compared to about half of the untreated men—a finding that is at least partially confirmed in our multivariate analyses controlling for demographics, relationship status, and prior behavior. However, compliance with the mental health referral and batterer program completion are inter-related. Men who dropout of the batterer program are much less likely to bother with mental health treatment.

The supplemental treatment was also associated with lower rates of re-assault. Those men who completed an evaluation were nearly a third less likely to re-assault their partner during the full 12-month follow-up (22% vs. 31%; n.s.; n=119), and a third less likely to re-assault if they received treatment (19% vs. 30%; n.s.; n=119). The very small sample of men receiving treatment contributes to these differences not being statistically significant and to the likelihood that unmeasured characteristics, such as personal motivation or perceived incentives, are influencing the results.

## IMPLICATIONS

The extra efforts, procedures, coordination, and costs to implement a mandatory referral system did not improve outcomes overall. This result is, no doubt, related to such a small portion of men actually receiving treatment and the consequential lack of improvement for the referred men as a whole. On the other hand, the majority of men who do receive treatment appear to appreciate and benefit from it. On the whole, it is difficult to justify the apparatus of the mandatory referral for this relatively small portion of men and the small net improvement in outcomes. The option of referral for men with observed or self-reported problems seems justified however. And case-management is likely to facilitate and increase the compliance to such referrals.

In any case, we are left with several unanswered speculations. Would a fuller mandate, namely stricter court oversight with sanctions for non-compliance, substantially increase the portion of men in treatment and improve outcomes? Or would the non-compliant men continue to resist compliance, or not cooperate if they did appear for an evaluation? We can also only speculate on the contribution of supplemental mental health treatment to batterer program outcomes. Can such treatment adequately accommodate court-mandated clients; does such treatment substantially impact court-mandated clients; would treatment significantly reduce abusive and violent behavior?

One instructive finding in response to the call for mental health assessment and referral in the field is that batterer program participants are not likely to comply with voluntary referrals. Some degree of accountability in terms of court-mandate, oversight and sanctions will substantially improve compliance, and the men who do comply under even modest coercion will benefit from it. This assertion echoes the general consensus in the drug and alcohol treatment field based on meta-analyses of so-called “drug courts”: coerced treatment is more effective in terms of compliance and outcomes than voluntary treatment (e.g., Wilson, Mitchell, & MacKenzie, 2006). As discussed in the



section above on “Implementation,” the question remains about how to establish a consistent and integrated referral system in an efficient and cost effective manner.

### **LIMITATIONS AND FUTURE RESEARCH**

The greatest strength of our research project is also its greatest weakness—that is, its lack of experimental conditions. A controlled clinical trial or experimental evaluation of supplemental mental health treatment would, no doubt, produce a clearer picture of the impact of such treatment over other options. It would however still not address the fundamental issue of dose-response—that is, whether coercing men into supplemental treatment improves outcomes. It also would neglect the “real world” circumstances and context of such treatment that the current study has exposed. The current study also failed to address the outcomes of a fully implemented mandatory referral—one in which the courts did consistently monitor and sanction non-compliance. Consequently, we do not know whether the ideal of full compliance is worth working towards and would make a difference.

Another limitation is the singular form of screening used in the research project. While there is some evidence that the BSI produces similar results to a more comprehensive as well as a shorter screening instrument, other screening and evaluation approaches may reduce over-screening and lead to more efficient and compliant referral. Individual evaluations at batterer program intake might also be tested against various screening instruments and protocol. The more individual evaluations might produce a more nuanced assessment of the men’s problems, needs, and motivation.

Other approaches to treatment might also be evaluated. Approaches that integrate domestic violence education with mental health treatment, or include mental health treatment as an auxiliary component of the batterer program, would reduce the complexity of referral and likely increase compliance. Whether such approaches would also reduce re-assault, or merely diffuse the domestic

violence education remains to be seen. Other more comprehensive approaches might be explored as well. For instance, a small residential program is working with problematic batterers in Perth, Australia. It draws on the principles of milieu therapy to address the variety of the men's problems, and, at the same time, professes a focus on the gender-based issues of domestic violence.

Finally, there is the uncertain assumption that mental health treatment might help reduce domestic violence. The relationship of mental health problems to violent behavior is, no doubt, an inconsistent and complex one. What specific problems or diagnoses should be of concern, and which of those are responsive to treatment? This sort of question might best be investigated using a more purposive sampling. That is, men with established diagnoses in a mental health clinic and screened for domestic violence might be randomly assigned to domestic violence education in addition to their mental health treatment. A comparison of their outcomes to men receiving only mental health treatment or only domestic violence education could produce some revealing results.

In sum, our research project raises several questions as well as answers some unintended ones. It especially exposes the difficulties and shortcomings in mandated referral for mental health treatment, as well as illustrates means to improve treatment compliance and outcomes for the compliant men. It appears to confirm the challenges facing the movement toward "coordinated community responses," while showing that increased oversight and accountability increases compliance to some degree. It also reinforces the increased attention to psychopathology in the field of domestic violence. A substantial portion of men exhibiting mental health problems were identified, but the impact of receiving mental health treatment remains unclear. More research is needed, at the one end, on improving implementation of coordinated or more comprehensive intervention, and, at the other end, on the relationship of mental health problems to domestic violence itself.

One member of our advisory panel observed that perhaps the most important aspect of our research project and ones like it is the discussion, thought and debate that it raises. Rather than a definitive result, it furthers our consideration and attention to domestic violence, and in the process helps to change community norms, attitudes, and practices. It is contributing to an on-going shift in consciousness and outlook. An encouraging and hopeful thought, but one that eludes research substantiation—at least for the moment.

## REFERENCES

- Allen, N. (2006). An examination of the effectiveness of domestic violence coordinating councils. *Violence Against Women*, 12, 46-67.
- Austin, J., & Dankwort, J. (1999). Standards for batterer programs. *Journal of Interpersonal Violence*, 14, 152-169.
- Bancroft, L. (2002). *Why does he do that? Inside the minds of angry and controlling men*. New York: Putnam.
- Bennett, L., Stoops, C., Call, C., & Flett, H. (2007). Program completion and re-arrest in a batterer intervention system. *Research on Social Work Practice*, 17, 42-54.
- Bland, R., & Orn, H. (1986). Family violence and psychiatric disorder. *Canadian Journal of Psychiatry*, 31, 129-137.
- Boulet, J., & Boss, M. (1991). Reliability and validity of the Brief Symptom Inventory. *Psychological Assessment*, 3, 433-437.
- California State Auditor (2006, November). *Batterer intervention programs: County probation departments, the courts and program compliance*. Bureau of State Audits, Sacramento, CA.
- Choca, J., & Van Denburg, E. (1997). *Interpretive guide to the Millon Clinical Multiaxial Inventory, Second Edition*. Washington, DC: American Psychological Association.
- Cohen, J. (1994). The Earth is Round ( $p < .05$ ). *American Psychologist*, 49, 997-1003.
- Daly, J. E., & Pelowski, S. (2000). Predictors of dropout among men who batter: A review of studies with implications for research and practice. *Violence and Victims*, 15, 137-160.
- Derogatis, L. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine*, 13, 595-605.

- Derogatis, L. (1993). *BSI: Brief Symptom Inventory, administration, scoring, and procedures manual*. Baltimore, MD: Clinical Psychometric Research.
- Dobash, R.E., & Dobash, R.P. (2000). Evaluating criminal justice interventions for domestic violence. *Crime and Delinquency*, 46, 252–271.
- Dutton, D. (1998). *The abusive personality: Violence and control in intimate relationships*. New York: Guilford.
- Dutton, D., & Corvo, K. (2006). Transforming a flawed policy: A call to revive psychology and science in domestic violence research and practice. *Aggression and Violent Behavior*, 11, 457-483
- Dutton, D., & Kropp, P. (2000). A review of domestic violence risk instruments. *Trauma, Violence & Abuse*, 1, 171-181.
- Efron, B. & Feldman, D. (1991). Compliance as an explanatory variable in clinical trials. *Journal of the American Statistical Association*, 86, 9-17.
- Ehrensaft, M., Moffitt, T., & Caspi, A. (2004). Clinically abusive relationships in an unselected birth cohort: Men's and women's participation and developmental antecedents. *Journal of Abnormal Psychology*, 113(2), 258-270.
- Gamache, D. J., Edleson, J. L., & Schock, M. D. (1988). Coordinated police, judicial, and social service response to woman battering: A multiple-baseline evaluation across three communities. In G. T. Hotaling, D. Finkelhor, J. T. Kirkpatrick, & M. A. Straus (Eds.), *Coping with family violence: Research and policy perspectives* (pp. 193-209). Newbury Park, CA: Sage.

- Gavazzi, S., Julian, T., & McKenry, P. (1996). Utilization of the Brief Symptom Inventory to discriminate between violent and nonviolent male relationship partners. *Psychological Reports, 79*, 1047-1056.
- Gondolf, E. (1997a). Patterns of re-assault in batterer programs. *Violence and Victims, 12*, 373-387.
- Gondolf, E. (1997b, May). An experimental evaluation of child abuse prevention classes for court-ordered women batterers. Final report submitted to the Child Trust Fund of Pennsylvania, Pennsylvania Department of Welfare, Harrisburg, PA.
- Gondolf, E. (1998). *Assessing woman battering in mental health services*. Thousand Oaks, CA: Sage Publications.
- Gondolf, E. (1999a). Characteristics of court-mandated batterers in four cities: Diversity and dichotomies. *Violence Against Women, 5*, 1277-1293.
- Gondolf, E. (1999b). MCMI results for batterer program participants in four cities: Less “pathological” than expected. *Journal of Family Violence, 14*, 1-17.
- Gondolf, E. (1999c). A comparison of re-assault rates in four batterer programs: Do court referral, program length and services matter? *Journal of Interpersonal Violence, 14*, 41-61.
- Gondolf, E. (2000a). Human subject issues in batterer program evaluation. *Journal of Aggression, Maltreatment, and Trauma, 4*(1), 273-297.
- Gondolf, E. (2000b). Mandatory court review and batterer program compliance. *Journal of Interpersonal Violence, 15*(4), 438-437.
- Gondolf, E. (2000c). Re-assault at 30-months after batterer program intake. *International Journal of Offender Therapy and Comparative Criminology, 44*, 111-128.
- Gondolf, E. (2002). *Batterer intervention systems: Issues, outcomes, and recommendations*. Thousand Oaks, CA: Sage Publications.

- Gondolf, E. (2006). Cautions about applying neuroscience to batterer intervention. *Juvenile and Family Justice Today*, 16(1), 20-22.
- Gondolf, E. (2007). Culturally-focused batterer counseling for African American men: A clinical trial of re-assault and re-arrest outcomes. *Criminology and Public Policy*, 6, 341-366.
- Gondolf, E. (in press-a). Implementation of casemanagement for batterer program participants. *Violence Against Women*.
- Gondolf, E. (in press-b). Theoretical and research support for the Duluth Model: A reply to Dutton and Corvo. *Aggression and Violent Behavior*.
- Gondolf, E., & Fisher, F. (1988). *Battered women as survivors: An alternative to treating learned helplessness*. Lexington, MA: Lexington Books.
- Gondolf, E., & Jones, A. (2001). The program effect of batterer programs in three cities. *Violence and Victims*, 16(6), 693-704.
- Guba, E., & Lincoln, Y. (1989). *Fourth generation evaluation*. Thousand Oaks, CA: Sage Publications.
- Hamberger, L. K., (2002). The men's group program: A community-based, cognitive-behavioral, pro-feminist intervention program (pp. 7-1—7-43). In E. Aldarondo & F. Mederos (Eds.), *Men who batter: Intervention and prevention strategies in a diverse society*. New York: Civic Research Institute.
- Hamberger, L., & Hastings, J. (1991). Personality correlates of men who batter and nonviolent men: Some continuities and discontinuities. *Journal of Family Violence*, 6, 131-148.
- 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. *Violence and Victims*, 11, 277-292.

- Heckert, A., & Gondolf, E. (2004). Battered women's perceptions of risk versus risk factors and instruments in predicting repeat re-assault. *Journal of Interpersonal Violence*, 19 (7), 778-800.
- Heckert, A., & Gondolf, E. (2005). Do multiple outcomes and conditional factors improve prediction of domestic violence? *Violence and Victims*, 20, 3-24.
- Heckert, A., Matula, D., & Gondolf, E. (2000). Women's accounts of domestic violence. *Women and Criminal Justice*, 12, 95-120.
- Holtzworth-Munroe, A., & Stuart, G. (1994). Typologies of male batterers: Three subtypes and the differences among them. *Psychological Bulletin*, 116, 476-497.
- Jacobson, N., & Gottman, J. (1998). *When men batter women: New insights into ending abusive relationships*. New York: Simon and Shuster.
- Jones, A., & Gondolf, E. (2001). Time-varying risk factors for re-assault by batterer program participants. *Journal of Family Violence*, 16, 345-359.
- Kantor, G., & Straus, M. (1987). The "drunken bum" theory of wife beating. *Social Problems*, 34, 214-230.
- Kessler, R., Coccaro, E., Fava, M. (2007). Intermittent explosive disorder is common, has an early age of onset and is associated with the development of other mental disorders in the US population. *Evidence-Based Mental Health*, 32-32.
- Kessler, R., Demler, O., Frank, R., Olfson, M., Pincus, H., Walters, E., Wang, P., Wells, K., & Zaslavsky, A. (2005). Prevalence and treatment of mental disorders, 1990 to 2003. *New England Journal of Medicine*, 352, 2515-2523.



- Labriola, M., Rempel, M., O'Sullivan, C., & Frank, P. (2007, March) Court responses to batterer program noncompliance: A national perspective. Report submitted to the National Institute of Justice, Washington, DC.
- Marshall, L. (1992). Development of the severity of violence against women scales. *Journal of Family Violence*, 7, 103-121.
- Mederos, F. (2002). Changing our visions of intervention: The evolution of programs for physically abusive men (pp. 1-2—1-26). In E. Aldarondo & F. Mederos (Eds.), *Men who batter: Intervention and prevention strategies in a diverse society*. New York: Civic Research Institute.
- Millon, T. (1994). *Millon Clinical Multiaxial Inventory-III manual*. Minneapolis, MN: National Computer Systems.
- Moffitt, T. E., Caspi, A., Rutter, M., & Silva, P. A. (2001). *Sex differences in antisocial behavior*. Cambridge: Cambridge University Press.
- Monahan, J., Steadman, H., Sliver, E., Appelbaum, P., Robbins, P., Mulvey, E., Roth, L., Grisso, T., & Banks, S. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York: Oxford University Press.
- Morey, L. (1997). *Personality Assessment Screener*. Lutz, FL: Psychological Assessment Resources.
- Murphy, C. M., Musser, P. H. & Maton, K. I. (1998). Coordinated community intervention for domestic abusers: Intervention system involvement and criminal recidivism. *Journal of Family Violence*, 13, 263-284.
- Murray, K. (2001). *Validating the Brief Symptom Inventory for use with a correctional population*. Dissertation Abstracts International, Section A: Humanities & Social Sciences, 61, 2607.

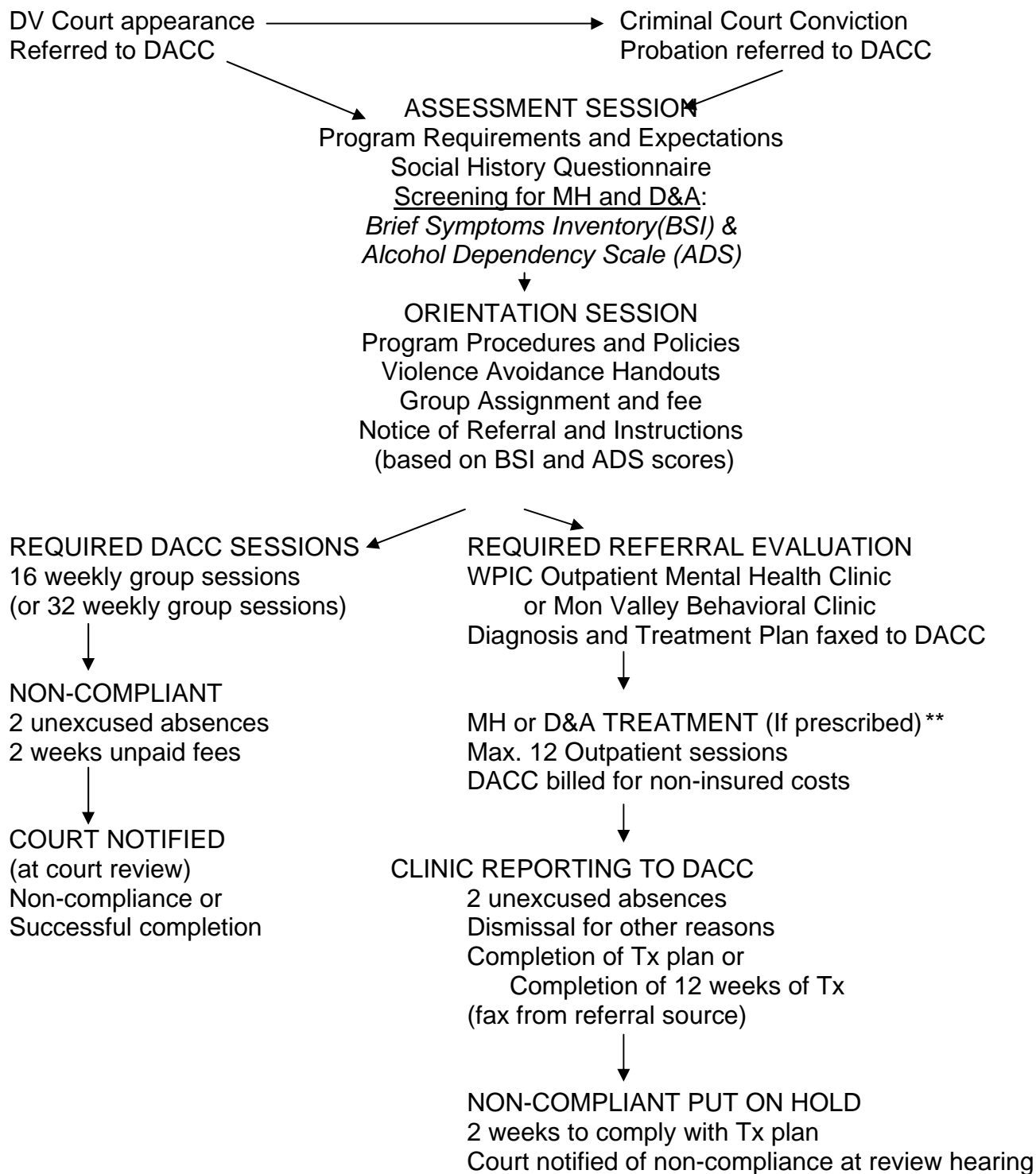
- O’Leary D. (1993). Through a psychological lens: Personality traits, personality disorders, and levels of violence. In R. Gelles & D. Loseke (Eds.), *Current controversies on family violence*. Thousand Oaks, CA: Sage Publications.
- Pawson, R., & Tilly, N. (1997). *Realistic evaluation*. Thousand Oaks, CA: Sage Publications.
- Pence, E. & Paymar, M. (1993). *Education groups for men who batter: The Duluth model*. New York: Springer.
- Pence, E., & McDonnell, C. (1999). Developing policies and protocols. In E. Pence & M. Shepard (Eds.), *Coordinated community response to domestic violence*. Thousand Oaks, CA; Sage.
- Pence, E., & Shepard, M. (Eds.). (1999). *Coordinated community response to domestic violence*. Thousand Oaks, CA; Sage.
- Peterson, C. (1989). Review of the Brief Symptom Inventory. In J. Conoley & J. Kramer (Ed.), *The tenth mental measurements yearbook*. Lincoln, NE: Buros Institute of Mental Measurements.
- Pettit, L., & Smith, R. (2002). The AMEND model (pp. 8-1—8-23). In E. Aldarondo & F. Mederos (Eds.), *Men who batter: Intervention and prevention strategies in a diverse society*. New York: Civic Research Institute.
- Piersma, H., Boes, J., & Reaume, W. (1994). Unidimensionality of the brief symptom inventory (BSI) in adult and adolescent inpatients. *Journal of Personality Assessment*, 63, 338-344.
- Project MATCH Research Group (1997). Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcome. *Journal of Studies on Alcohol*, 58, 7-29.
- Rice, M. (1997). Violent offender research and implications for the criminal justice system. *American Psychologist*, 52, 414-423.
- Roehl, J., & Guertin, K. (2000). Intimate partner violence: The current use of risk assessments in sentencing offenders. *Justice System Journal*, 21, 171-197.

- Russell, M. (1995). *Confronting abusive beliefs: Group treatment for abusive men*. Thousand Oaks, CA: Sage.
- Selzer, M. (1971). The Michigan Alcoholism Screening Test (MAST): The quest for a new diagnostic instrument. *American Journal of Psychiatry*, 3, 176-181.
- Shea, S. (1988). *Psychiatric interviewing: The art of understanding*. Philadelphia, PA: W.B. Saunders.
- Shepard., M., & Campbell, J. (1992). The Abusive Behavior Inventory: A measure of psychological and physical abuse. *Journal of Interpersonal Violence*, 7, 291-305.
- Sherman, L. (1992). *Policing domestic violence: Experiments and dilemmas*. New York: Free Press.
- Skinner, H., & Horn, J. (1984). *Alcohol Dependence Scale (ADS) user's guide*. Toronto, Canada: Additional Research Foundation.
- Stordeur, R., & Stille, R. (1989). *Ending men's violence against their partners*. Thousand Oaks, CA: Sage.
- Stosny, S. (1995). *Treating attachment abuse: A compassion approach*. New York: Springer.
- Straus, M. (1979). Measuring intrafamily conflict and violence: The Conflict Tactics (CT) scales. *Journal of Marriage and the Family*, 41, 75-88.
- Swanson, J. (1994). Mental disorder, substance abuse, and community violence: An epidemiological approach (pp. 101-136). In J. Monahan & H. Steadman (Eds.), *Violence and mental disorders: Developments in risk assessment*. Chicago, IL: University of Chicago Press.
- Tjaden, P. G., & Thoennes, N. (2000). *Full Report of Prevalence, Incidence and Consequences of Violence Against Women: Findings from the National Violence Against Women Survey (NCJ 183781)*. U.S. Department of Justice, National Institute of Justice and Centers for Disease Control and Prevention Research Report, Washington, DC.

- Tolman, R. (1989). The development of a measure of psychological maltreatment of women by their male partners. *Violence and Victims*, 4, 159-177.
- Tolman, R., & Bennett, L. (1990). A review of quantitative research on men who batter. *Journal of Interpersonal Violence*, 5, 87-118.
- Visher, C., Newmark, L., & Harrell, A. (2006). Final report on the evaluation of the Judicial Oversight Demonstration (Volume 2): Findings and lessons on Implementation. Washington, DC: Urban Institute.
- White, R., & Gondolf, E. (2000). Implications of personality profiles for batterers treatment. *Journal of Interpersonal Violence*, 15, 467-488.
- Wilson, D., Mitchell, O., & MacKenzie, D. (2006). A systematic review of drug court effects on recidivism. *Journal of Experimental Criminology*, 2, 459-487.
- Yllo, K. (1993). Through a feminist lens: Gender, power, and violence. In R. Gelles & D. Loseke (Eds.), *Current controversies on family violence*. Thousand Oaks, CA: Sage Publications.
- Zimmerman, M. (2002). *Psychiatric Diagnostic Screening Questionnaire (PDSQ)*. Los Angeles, CA: Western Psychological Services.

## **FIGURES & TABLES**

**FIGURE 1: DACC MENTAL HEALTH ENHANCED REFERRAL**



NOTE: DACC representative is to report non-compliance to court during review hearing, and court is to sanction man with jailing, additional DACC sessions, fines, or full prosecution.

## FIGURE 2: STEPS FOR REQUIRED REFERRAL

Below are the STEPS for required referrals from DACC. Check each box as you complete that step and go on to the next one. More details are on the sheet “DACC Procedures for Required Referrals.”

### CALL THE CLINIC FOR AN APPOINTMENT.

See the “Referral List” for the phone number and call *with in the next 3 days*.

Tell who answers that “DACC has referred you for an evaluation.”

Clinic phone number:

---

---

### GO TO YOUR APPOINTMENT—DON'T MISS IT!

Give the receptionist the “Clinical Referral Instructions” and the two forms attached to it (YELLOW).

Tell the receptionist that you need to sign a release form.

Appointment Date and Time:

---

---

### ATTEND THE SESSIONS INDICATED IN YOUR TREATMENT PLAN

You are required to do the clinic treatment plan in addition to your DACC groups for a maximum of 3 months (12 weeks).

The clinic will notify DACC by fax when you complete the treatment plan.

Sessions Day and Time:

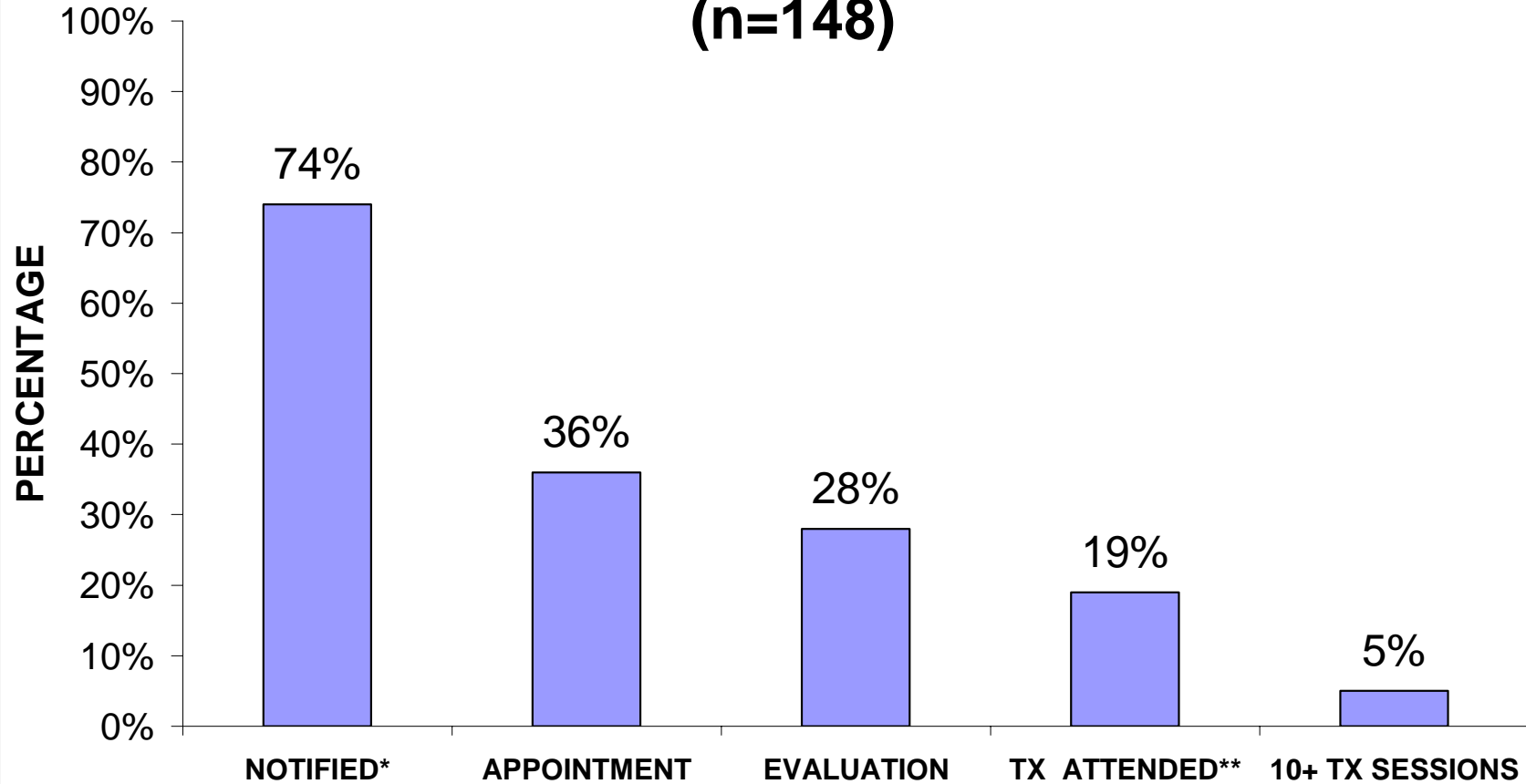
\_\_\_\_\_

NOTE: Let the DACC main office know if you have any problems with making an appointment or completing your treatment plan (~~XXXXXXX~~). If you should get a bill, bring it or send it to the DACC main office.

Your Name: \_\_\_\_\_

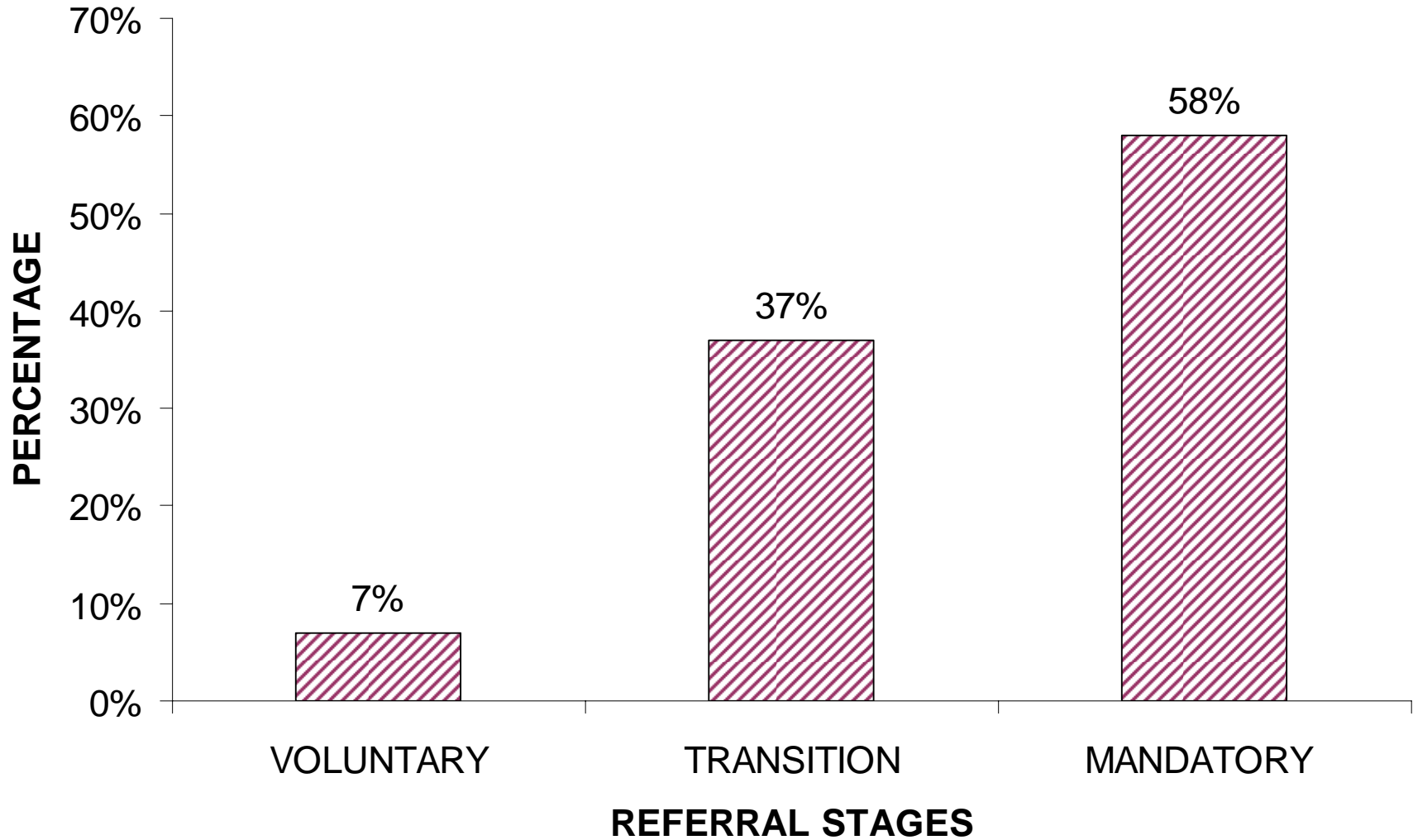


### FIGURE 3: COMPLIANCE WITH MANDATED REFERRAL BASED ON CLINICAL RECORDS (n=148)



\*based on Men's 5-mo. Follow-up; \*\*8% of sample already in treatment at batterer at time of evaluation leaving 10% starting treatment as a result of referral

**FIGURE 4: APPOINTMENT FOR CLINICAL EVALUATION REPORTED BY REFERRED BATTERER PROGRAM PARTICIPANTS (based on Men's 5-mo. follow-up; n=254)**



## **TABLE 1: COMPONENTS FOR MENTAL HEALTH REFERRAL**

### **Court Supervision (Domestic Violence Court)**

Arrested perpetrators into preliminary hearing within week  
Referred to batterer program as condition of bond  
Mental health referral part of batterer program compliance  
Two court reviews of compliance (payment and attendance)  
Sanctions for non-compliance (fines, jail, prosecution, more sessions)

### **Domestic Violence Counseling (Batterer Program)**

14 weekly group sessions  
Cognitive-behavioral approach  
Group leaders with BA and MA degrees  
Multiple sites in city  
Fees on sliding scale  
Screened with BSI and ADS at intake  
Notified at orientation of referral  
Liaison presents compliance to court

### **Supplemental Mental Health Treatment (Mental Health Clinics)**

Referred for evaluation to two clinics  
Required to attend prescribed treatment  
Evaluation and treatment at no-cost  
Batterer program notified of compliance

## **TABLE 2: CHALLENGES IN IMPLEMENTING MANDATED MENTAL HEALTH REFERRAL**

**Assessment**—consistent screening at four intake sites

**Notification**—consistent notification of assessment results and referral procedures

**Mental Health Evaluation**—establishing a central contact and timely response

**Billing**—following procedures for insurance coverage and an alternative coverage

**Treatment Prescribed**—obtaining a diagnosis that qualifies for treatment.

**Verification of Evaluation and Treatment**—getting documentation of compliance

**Informing the Court of Compliance**—need for batterer program to report and assert non-compliance.

**Court Response to Non-Compliance**—issuing sanctions for non-compliance

## TABLE 3: REFERRAL SAMPLE RECRUITMENT AND ELIGIBILITY

### SAMPLE RECRUITMENT

#### 1043 screened at program intake 6/04-4/06

12 deleted for refusal, 34 not applicable, or 85 invalid BSI scores=131 (13% of intake are deleted; 912 total undeleted)

#### 515 eligible for referral (BSI or ADS positive)

26 deleted for not applicable, 14 initial incident not partner but previous violence with a partner; 6 no partner information, 7 previously admitted to study, 4 underage, 8 not properly notified, 4 refused outright, 4 reason uncertain =73 deleted (14% of eligible are deleted; 442 total undeleted)

#### 479 referred and eligible for follow-up study

(46% of total 1043 screened at intake)

BSI positive=426 (41% of total screened)

ADS positive only (i.e., not BSI positive)=15

Judge-mandated only (i.e., not BSI or ADS positive)=37

NOTE: Total subjects with *invalid* BSI=95 (9% of total screened): 85 not eligible for referral, 8 referred by judges, and 3 referred because of positive ADS score.

### REFERRAL ELIGIBILITY AND STAGE

#### Combined eligibility

ADS+BSI positive=31

Judge-mandated with BSI or ADS positive=36

Any ADS positive=46 (9% of total 479 referred men)

Any judge-mandated=73 (15% total 479 referred men)

#### Referral Stages

Voluntary referral 6/1/04-1/15/05 (38%)= 182

Transitional referral 1/15/05-8/31/05 (35%)=166

Mandated referral 9/1/05-4/15/06 (27%)= 131

**TABLE 4: BSI POSITIVE SUBSCALES FOR REFERRAL SAMPLE**

<b>BSI subscale</b>	<b>Percent</b>
Somatization	31%
Obsessive-Compulsive	46%
Interpersonal Sensitivity	51%
<b>Depression</b>	<b>59%</b>
Panic Anxiety	41%
Hostility	34%
Phobic Anxiety	34%
<b>Paranoid Ideation</b>	<b>72%</b>
<b>Psychoticism</b>	<b>57%</b>
<b>Global Score Index</b>	<b>58%</b>
Any Subscale=80 (maximum score)	30%
GSI=80 (maximum score)	15%
BSI Subscales Positive: 6 or more of 9 subscales	33%
Alcohol Dependence Scale (>13)	9%

---

N=479

**TABLE 5: SELECTED CHARACTERISTICS OF PROGRAM PARTICIPANTS (percentages)**

<b>Characteristics</b>	<b>BSI Result</b>		<b>Total (N=910)</b>
	<b>Negative (n=431)</b>	<b>Positive (n=479)</b>	
<b>DEMOGRAPHICS</b>			
< 30 years old	39	44	41
Race (Non-white/ Minority)	51	50	51
Some college	29	31	30
Unemployed (or partial employment)	41	50	45**
<b>RELATIONSHIP STATUS</b>			
Married	28	26	27
Living with partner	44	45	44
See Partner Daily	50	50	50
Children Living with	52	41	46***
<b>PAST BEHAVIOR</b>			
Parent hit parent	14	31	23***
Used drugs (past year)	13	26	19***
Drunk monthly+ (past year)	17	31	24***
Alcohol Dependence (Ads>13)	0	10	5***
Multiple Symptom Areas (>5 positive BSI subscales)	0	33	19***
<b>DOMESTIC VIOLENCE</b>			
Threats (past 3 mos.)	19	28	21***
Severe Assaults (ever)	7	16	12***
Bruised (current incident)	23	34	29***
Partner also arrested (current incident)	9	11	10
<b>PAST INTERVENTION</b>			
Protection Order (ever)	22	23	22
DV arrest (ever)	19	27	23**
Batterer Counseling (ever)	17	18	18
Non-DV arrest (ever)	42	49	45*
Probation/parole (current)	27	24	25

**MAN'S PERCEPTIONS**

Very unlikely to hit again	98	89	91*
Partner feels very safe	84	72	77***
Program attendance (3 or more sessions)	82	83	83
Program completion (16 or more sessions)	62	57	58

**MENTAL HEALTH**

Any psychological treatment (ever)	14	34	24***
Psychological counseling (ever)	9	25	17***
Psychiatric hospital or outpatient Tx (ever)	4	12	8***
Prescription medication (current)	4	12	8***

**BENEFIT FROM ADDITIONAL TREATMENT  
(some or great extent)**

Batterer counseling	46	56	42*
Mental Health Tx	17	34	26***
Alcohol and/or drug Tx	15	32	24***

---

\*p<.05; \*\*p<.01; \*\*\*p<.001; N=910  
 Note: DV= domestic violence; Tx= treatment



**TABLE 6: REFERRAL STAGES FOR BATTERER PROGRAM PARTICIPANTS (PERCENTAGE)**

Characteristics	Referral Stage			Total (N= 479)
	Voluntary (n= 182)	Transitional (n= 166)	Mandatory (n= 131)	
<b>DEMOGRAPHICS</b>				
< 30 years old	44	44	44	44
Some college	32	33	26	31
Unemployed (other)	47	51	53	50
<b>RELATIONSHIP STATUS</b>				
Married	23	24	31	26
Living with partner	44	42	50	45
See Partner Daily	47	52	50	50
Children Living with	36	45	43	41
<b>PAST BEHAVIOR</b>				
Parent hit parent	35	28	30	31
Used drugs (past year)	27	23	27	26
Drunk monthly+ (past year)	34	28	30	31
Alcohol Dependence (Ads>13)	13	7	10	10
Multiple Symptom Areas (6 or more positive BSI subscales)	36	36	25	<b>33*</b>
<b>DOMESTIC VIOLENCE</b>				
Threats (past 3 mos.)	33	28	23	28
Severe Assaults (ever)	16	16	17	16
Bruised (current incident)	35	33	33	34
Partner also arrested (current incident)	11	15	7	11
<b>PAST INTERVENTION</b>				
Protection Order (ever)	25	20	23	23
DV arrest (ever)	27	29	24	27
Batterer Counseling (ever)	18	23	12	<b>18*</b>
Non-DV arrest (ever)	54	45	45	49

Probation/parole (current)	25	25	20	24
----------------------------	----	----	----	----

**PERCEPTIONS**

Very unlikely to hit again	11	11	13	12
Partner feels very safe	69	72	75	72
Program attendance (3 or more sessions)	80	83	89	83
Program completion (16 or more sessions)	57	60	52	57

---

\*p<.05

**TABLE 7: DIAGNOSIS BASED ON CLINICAL RECORDS (n=38)**

Clinical Diagnoses	Percent	Number
Adjustment disorder	37%	14
Dual diagnoses (with drug/alcohol Dx)	21%	8
Drug or alcohol dependence (only)	16%	6
Depression	13%	5
Impulse disorder	8%	3
Deferred	5%	2

**TABLE 8: AGREEMENT OF POSITIVE BSI SUBSCALES WITH CLINICAL DIAGNOSES (percent and number of total diagnosed cases, n=38)**

Clinical Diagnosis	BSI Subscale									GSI*
	Somatization	Obsessive Compulsive	Interpersonal Sensitivity	Depression	Panic Anxiety	Hostility	Phobic Anxiety	Paranoid Ideation	Psychoticism	
Adjustment Disorder	14%(5)	17%(6)	19%(7)	22%(8)	16%(6)	11%(4)	11%(4)	31%(11)	19%(7)	19%(7)
Dual Diagnosis	6%(2)	8%(3)	11%(4)	8%(3)	8%(3)	8%(3)	3%(1)	14%(5)	11%(4)	8%(3)
Substance Abuse	3%(1)	0%(0)	3%(1)	8%(3)	6%(2)	3%(1)	0%(0)	3%(1)	6%(2)	3%(1)
Depression Disorder	11%(4)	8%(3)	8%(3)	11%(4)	8%(3)	6%(2)	8%(3)	8%(3)	11%(4)	8%(3)
Impulse Control	3%(1)	3%(1)	6%(2)	6%(2)	3%(1)	3%(1)	3%(1)	6%(2)	6%(2)	6%(2)
Deferred Diagnosis	0%(0)	0%(0)	0%(0)	0%(0)	0%(0)	0%(0)	0%(0)	0%(0)	0%(0)	0%(0)

\* Global Severity Index

**TABLE 9: SCREENING INSTRUMENTS COMPARED FOR CONCURRENT VALIDITY**

<b>Instrument</b>	<b>Number of Items</b>	<b>Focus of Items</b>	<b>Time Frame</b>	<b>Scoring Levels</b>
BSI (Brief Symptom Inventory)	53	Distress	Previous week	2
PDSQ (Psychiatric Diagnostic Screening Questionnaire)	111	Psychiatric symptoms	Past 6 mos.	6
PSA (Personality Assessment Screener)	22	Problematic Traits	Current	5

**TABLE 10: MEN’S RESPONSE TO MENTAL HEALTH REFERRAL (percentage)**

Topic	Case-Management Interview		Men’s 5-mo. Follow-up		
	Transitional Stage (n=21)	Mandatory Stage (n=95)	Voluntary Stage (n=95)	Transitional Stage (n=74)	Mandatory Stage (n=75)
Problems contacting MH clinic	11	30	0	9	15
Problems with visiting MH clinic	0	14	NA	8	8
Clinic Contact					
Made appointment	38	50	7	37	58
Obtained evaluation	14	27	5	37	55
TX recommended	4	13	NA	NA	NA
TX received	9	4		17	24
10+ Tx sessions	0	25	3	7	17

**TABLE 11: MEN’S RESPONSE TO MENTAL HEALTH TREATMENT NEEDS (percentage)**

Topic	Case-Management Interview		Men’s 5-mo. Follow-up		
	Transitional Stage (n=21)	Mandatory Stage (n=95)	Voluntary Stage (n=95)	Transitional Stage (n=74)	Mandatory Stage (n=75)
Need MH or D&A help	43	36	24	27	24
Symptoms					
None	52	57	73	73	73
Anger	19	4	15	21	19
Depressed	10	5	19	17	22
Anxious	0	7	17	17	19
Fearful/distrustful	0	1	7	13	9
Confused	0	1	12	10	6
Stressed/distressed	2	16	25	27	25
Alcohol problems	5	11	9	15	9
Drug problems	5	2	6	9	3
Other	5	4	0	0	3
Other assistance needed	NA	NA	30	38	35

**TABLE 12: OUTCOMES FOR REFERRAL STAGES (PERCENTAGES)**

Outcomes	Referral Stage			Total (n=479)
	Voluntary Referral (n=182)	Transition Referral (n=166)	Mandated Referral (n=131)	
<b>PROGRAM COMPLETION</b> (n=479)				
3 or more weeks of program	80	83	88	83
Completion (16 sessions)	57	60	52	57
<b>RESSAULT</b>				
0-6 mos. (n=344 respondents)	30	21	23	19
0-12 mos. (n=308 respondents)	33	32	30	32
3-12 months (n=308 respondents)	25	25	26	25
<b>OTHER ABUSE</b> (n=308)				
Threats 0-12 mos.	36	47	43	42
Severe re-assault 0-12 mos.	16	12	16	15
Bruised or injured 0-12 mos.	18	19	18	18
<b>WOMEN'S PERCEPTIONS</b> (n=292)				
Overall well-being at 12 mos.	76	66	77	73
Feel "very safe" at 12 mos.	73	69	54	66*
"Very unlikely" to be hit at 12 mos.	67	66	50	62*
<b>NO PARTNER CONTACT</b> at 12 mos. (n=308)				
	29	23	12	22*

\*p<.05

**TABLE 13: OUTCOMES FOR COMPLIANCE TO MENTAL HEALTH REFERRAL  
(Percentages for compliance; n=148)**

Outcomes	Evaluation Obtained		Treatment Received	
	No (n=107)	Yes (n=48)	No (n=120)	Yes (n=28)
<b>PROGRAM COMPLETION</b> (n=148)				
3 or more weeks of program	83	98**	85	96
Completion (16 sessions)	38	76*	45	68**
<b>RESSAULT</b> (n=118)				
0-6 mos.	20	6	19	8
0-12 mos.	31	22	30	19
3-12 months	29	15	28	17
<b>OTHER ABUSE</b> (n=118)				
Threats 0-12 mos.	50	26	42	48
Severe re-assault 0-12 mos.	18	9	16	13
Bruised or injured 0-12 mos.	22	4	22	4
<b>WOMEN'S PERCEPTIONS</b> (n=118)				
Overall well-being at 12 mos.	48	60	53	45
Feel "very safe" at 12 mos.	55	65	60	48
"Very unlikely" to be hit at 12 mos.	77	74	71	90
<b>NO PARTNER CONTACT</b> at 12 mos. (n=118)	18	9	18	8

\*p<.001; \*\*p<.05

Note: Respondent sample based on *any* report during follow-up period.

**TABLE 14: RE-ARRESTS FOR REFERRAL STAGES (PERCENTAGES)**

Category of Crime	Referral Stage		Total (n=177)
	Voluntary Referral (n=86)	Mandated Referral (n=91)	
<b>ARREST PRIOR TO STUDY</b>			
Domestic violence	54	57	55
Other crimes	59	56	58
<b>ARREST DURING FOLLOW-UP</b>			
Domestic violence	7	4	6
Other violence	11	6	8
Domestic violence or other violence	16	9	12
Alcohol or drug related crime	9	8	9
Other crimes	12	3	8*
Any crime	29	18	23**
2 or more total crimes	9	6	7

\*p<.05; \*\*p<.1

**TABLE 15: RE-ARRESTS FOR NON-REFERRED VS. REFERRED TO MENTAL HEALTH CLINIC (PERCENTAGES)**

Category of Crime	Referral Status		Total (n=254)
	Not Referred (n=76)	Referred to Clinic (n=177)	
<b>ARREST PRIOR TO STUDY</b>			
Domestic violence	48	55	53
Other crimes	46	58	54**
<b>ARREST DURING FOLLOW-UP</b>			
Domestic violence	7	6	6
Other violence	14	8	10
Domestic violence or other violence	21	12	15
Alcohol or drug related crime	7	9	8
Other crimes	5	7	7
Any crime	29	23	25
2 or more total crimes	7	7	7

\*p<.05; \*\*p<.1