A Large Sample Investigation of Batterer Intervention Program Attrition: Evaluating the Impact of State Program Standards

Frederick P. Buttell and Michelle Mohr Carney

Research on Social Work Practice 2008; 18; 177
DOI: 10.1177/1049731508314277

A Large Sample Investigation of Batterer Intervention Program Attrition: Evaluating the Impact of State Program Standards

Frederick P. Buttell
Tulane University

Michelle Mohr Carney
University of Georgia

Objective: The purpose of this study was to (a) investigate differences in demographic and psychological variables between treatment completers and dropouts among abusive men entering a court-mandated treatment program and (b) evaluate the goodness of fit of a logistic regression model for predicting attrition developed on a 16-week program by applying that model to the 26-week program. Method: The study employed a secondary analysis of 1,702 men, 850 completers and 852 dropouts, randomly selected from a larger pool of 3,595 men in court-ordered treatment. Results: Analysis indicated that few of the demographic and psychological variables differentiated between completers and dropouts. A logistic regression model for the 26-week program was developed that correctly predicted treatment completion for 60% of the sample and that employed some of the same predictor variables as the 16-week model. Conclusion: Implications for predicting attrition using logistic regression models for enhancing retention rates were explored and discussed.

Keywords: *court-ordered batterers; domestic violence; spousal violence*

Data from separate national representative samples, the National Violence Against Women Survey (NVAWS) and the National Crime Victimization Survey (NCVS), suggest that intimate partner violence remains a serious social problem. The NVAWS indicated that an estimated 1.9 million women are physically assaulted annually in the United States and that violence against women is primarily intimate partner violence, with 64% of the women who were raped, physically assaulted, or stalked reporting that they were victimized by a current or former husband, cohabitating partner, boyfriend, or date, compared to only 16.2% of men (Tjaden & Thoennes, 2000). The data from the NCVS paint a similar picture, indicating that women were the victims of intimate partner violence about 5 times more often than males and that, on average, each year from 1993 to 1998, 22% of all U.S. female victims of violence were attacked by an intimate partner, compared to 3% of all male victims of violence (Rennison & Welchans, 2000). In terms of risk for injury to U.S. women from domestic violence, Kyriacou et al. (1999) have reported that the lifetime risk of severe injury as a result of domestic violence has been estimated to be 9% for women, with a lifetime risk of up to 22% for any type of injury from domestic violence.

Apart from the shelter movement, perhaps the most interesting and important societal response to intimate partner violence has been to begin to address the problem by working with the abuser. This response is predicated on the idea that any meaningful change effort will have to target that person in the relationship who initiates and maintains the problem—namely, the abusive partner or batterer. As part of the multifaceted effort to eradicate domestic violence in the United States, batterer intervention programs have been developed and implemented in conjunction with court programs (for a review, see Daly & Pelowski, 2000).

In brief, in the late 1980s, laws that mandated arrest in police calls involving domestic violence were enacted nationally, which served to dramatically increase the number of men arrested, prosecuted, and convicted of domestic violence offenses. In these cases, when the police respond to a domestic violence call and determine that a criminal domestic violence offense has occurred, the

Authors' Note: Correspondence may be addressed to Frederick P. Buttell, PhD, School of Social Work, Tulane University, 6823 St. Charles Avenue, New Orleans, LA 70118-5672; e-mail: buttell@tulane.edu.

Research on Social Work Practice, Vol. 18 No. 3, May 2008 177-188
DOI: 10.1177/1049731508314277

police assume the burden of pressing charges. The purpose of this legislation was to protect battered women because their violent partner was to understand that he would not be able to threaten or intimidate her into dropping charges. The net result of the legislation was that courts became inundated with domestic violence offenders. Consequently, because incarceration was not the option preferred by judges and court referral officers (i.e., space and time constraints), community-based batterer intervention programs (BIPs) proliferated as a sentencing option that judges could use in lieu of jail.

Since these programs came into existence, there have been dozens of single-site evaluation studies conducted on their effectiveness, along with several reviews and meta-analyses (e.g., Babcock, Green, & Robie, 2004; Davis & Taylor, 1999; Davis, Taylor, & Maxwell, 1998; Dunford, 2000; Feder & Forde, 2000; Feder & Wilson, 2005; Levesque, 1999; Morrison & Nesius, 2003). Although there is heated debate in the field regarding BIP effectiveness (e.g., Gondolf, 2001), one thing that is beyond debate is the fact that virtually all of the evaluation studies have suffered from very high attrition rates. In fact, because the BIP attrition rate nationally hovers between 50% and 75% (Daly & Pelowski, 2000), a number of studies have investigated demographic and psychological variables associated with BIP attrition. Unfortunately, the picture that emerges from this body of literature is more confusing than clarifying. For example, there is some literature that suggests that demographic variables (i.e., age, employment status, educational level, alcohol use, income, previous criminal history, and relationship status) can distinguish between treatment completers and dropouts, with dropouts tending to be younger, unemployed, less educated, more likely to abuse alcohol, and either single or separated and to have a previous criminal history (Buttell & Carney, 2002; Cadsky, Hanson, Crawford, & Lalonde, 1996; Chang & Saunders, 2002; Daly, Power, & Gondolf, 2001; DeMaris, 1989; Gruznski & Carrillo, 1988; Hamberger & Hastings, 1989; McCloskey, Sitaker, Grigsby, & Malloy, 2003; Scott, 2004). Other studies, however, have discovered either inconsistent or nonsignificant differences between treatment completers and dropouts on these variables (Buttell & Pike, 2002; Chen, Bersani, Myers, & Denton, 1989; DeHart, Kennerly, Burke, & Follingstad, 1999; DeMaris, 1989; Gruznski & Carrillo, 1988; Hamberger & Hastings, 1986, 1991; Hamberger, Lohr, & Gottlieb, 2000; Jones & Gondolf, 2002).

Recently, in an effort to make sense of this confusing body of literature, two articles have suggested that the existing BIP attrition studies may have been inadequately framed. Specifically, Buttell and Carney (2002) discovered that unexamined variables like judicial support for the program may interact with the variables examined in the study and confound generalizations to other programs and geographical locations. For example, in their study, one judge that referred to the program ensured that those men failing to attend the program would be met with immediate consequences (i.e., an arrest warrant would be immediately served by the police), whereas another judge that referred to the program would give men multiple chances to complete the program. They discovered that this variable was far more important in predicting dropout than either the demographic or psychological variables considered in their study. Consequently, because there is huge variability among treatment providers nationally on this construct and many others, Buttell and Carney (2004) argued that to create meaningful logistic regression models that accurately identify men at risk of premature dropout prospectively, every BIP needs to work in conjunction with researchers to develop models that capture their unique situation (e.g., population of batterers, judicial support for the program, etc.).

Despite the potential importance of this suggestion, ongoing changes in the field of BIPs make implementing it difficult. Specifically, in an effort to enhance the accountability of BIPs, states have passed legislation governing many aspects of BIP service provision. Research suggests that, at the time of this writing, 43 states have passed legislation mandating minimum standards for BIPs, and several more have legislation pending (State Standards, 2006). Among the many issues addressed by this legislation is program length. Most states have adopted legislation requiring BIPs to be a minimum of 26 weeks long, and some states (16%) require them to be 52 weeks long (Maiuro, Hagar, Lin, & Olson, 2001). This represents a significant change for most programs, which typically had 10- to 16-week programs prior to the implementation of the state standards. This lengthening of the BIP is particularly important to service providers seeking to understand the factors associated with attrition because the increased program length may interact with judicial support for the program and eliminate the usefulness of the regression models that were created using data involving the old program length. For example, in the era prior to program standards, a particular judge may have been willing to mandate a man to a 10- to 16-week BIP for verbal abuse (i.e., harassment) or mild physical violence (i.e., slapping) and ensure his compliance though the use of a bench warrant (discussed previously). However, that same judge may be willing to mandate that same man to 26 to 52 weeks of treatment but unwilling to

ensure compliance with a bench warrant, particularly if the man has completed some arbitrary number of sessions (e.g., 8 to 10 weeks). Currently, it is unclear whether the models used to predict attrition in BIPs prior to the implementation of state standards have any utility for predicting attrition in the longer programs because there is no research that directly addresses this question.

Given the confusing situation regarding BIP attrition in general and the absence of any studies investigating the utility of models developed on programs of shorter duration for predicting attrition in the new, longer state-legislated programs, the purpose of the present study was to both replicate and extend the work in this area by (a) attempting to create a predictive model that would assist the partnering BIP in correctly identifying men at greatest risk for dropping out of the program and (b) exploring the utility of a model developed from the 16-week program for predicting attrition in the new 26-week program. By focusing exclusively on men court ordered into treatment at the same program site, the potentially confounding effects of referral status (i.e., voluntary versus involuntary) and judicial support for the program could be controlled.

METHOD

Data Collection

This study employed a secondary analysis of data collected by the Domestic Abuse Center, a nonprofit agency in Columbia, South Carolina, that has been providing counseling services to men court ordered into treatment since 1982. The study was approved with exempt status through the University of South Carolina Institutional Review Board. At the intake interview, clients are required to complete a demographic questionnaire and four psychological instruments: Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1984), Spouse-Specific Assertiveness Scale (SSAS; O'Leary & Curley, 1986), Control of Partner Scale (CPS; Follingstad, Rutledge, McNeill-Harkins, & Polek, 1988), and the Propensity for Abusiveness Scale (PAS; Dutton, 1995). The agency uses this information to create a pretreatment assessment of the individual. At the conclusion of the treatment program, the agency requires that clients, once again, complete the instrument package.

Operational Definitions of *Treatment Completer* and *Dropout*

Given the high rates of attrition associated with batterer intervention programs discussed previously, many researchers have adopted a definition for the word *completer* that would be more accurately described as "partial completer" (Buttell & Carney, 2004). Specifically, because so many men fail to fully complete the batterer intervention program, researchers have conceptualized men as "completers" if they completed some fraction of the prescribed treatment (e.g., 4 weeks of a 12-week program). The problem with such a definition is that it equates men who complete a portion of the program with men who complete the entire program. Certainly, there are differences between these two groups, and the men who complete the entire program are likely to benefit more than men who complete a fraction of it (Hamberger & Hastings, 1988). If this is not the case, why have a treatment program that is 12 weeks long instead of one that is 4 weeks long? Clearly, the premise is that all of the material in the program is important for modifying behavior, and, therefore, an accurate definition of "completer" should entail finishing the entire program. Consequently, this study defined *treatment completers* as those men who successfully finished the 26-week treatment program and *dropouts* as those men who failed to successfully finish the program.

Participants

The agency has been using the combination of assessment tools described above since August 1999. At the time of this writing, 3,595 men have been referred to the 26-week program by the courts. Of this group of men, 1,941 men (54%) have dropped out or been terminated from the program, 899 men (25%) have graduated from the program, and 755 men (21%) are still active in the program. The sample for this study included all 26-week treatment completers ($n = 899$) and a random sample of 26-week treatment dropouts ($n = 900$). Within this sample of court-mandated referrals, 90% were referred following an arrest for domestic violence by domestic violence courts, probation, or parole officers, department of social services, or a diversion program. Of the 1,799 men originally selected for the sample, 25 were omitted from the analysis because they repeatedly gave the same rating to all the items and were clearly not discriminating among them or they failed to complete the entire instrument package. In addition, 72 were omitted from the analysis because of high scores (>13) on both the pre- and posttreatment BIDR Impression Management subscale. It was determined that their desire to give a favorable impression of themselves would likely negatively influence the accuracy of their responses. Thus, the final sample consisted of 850 male treatment completers and 852 male treatment dropouts in heterosexual relationships.

TABLE 1: Participant Characteristics

	Program Participants (N = 1,702)					
	Completers		Dropouts		Total	
Characteristic	%	*N*	%	*N*	%	*N*
Relationship status						
Married	61.5	523	57.6	491	42.4	25
Not married	38.5	327	42.4	361	57.6	34
Arrest prior to program**						
Yes	94.2	801	86.5	737	90.4	1,538
No	5.8	49	13.5	115	9.6	164
Race/ethnicity						
African American	50.6	430	51.2	436	50.9	866
Caucasian	49.4	420	48.8	416	49.1	836
Alcohol use						
Use at one time	85.8	729	87.3	744	88.3	1,473
Marijuana use*						
Use at one time	54.1	460	59.3	505	57.8	965
Prescription drug use						
Use at one time	41.9	356	41.7	355	41.8	711
	M	*SD*	*M*	*SD*	*M*	*SD*
Age**	35.89	9.2	33.74	9.0	34.81	9.2
Education (years)	12.28	2.1	12.17	2.2	12.22	2.1
Length of current relationship (months)*	94.58	82.0	85.10	74.2	90.08	78.5
Monthly income**	$2,234.44	$1,360.22	$1,908.01	$1,333.55	$2,080.51	$1,356.96

*$p < .05$. **$p < .001$.

Program completers and dropouts were compared using a variety of variables consistent with the batterer attrition literature previously discussed. Most participants were married (59.6%), remained involved in the relationship with the victim (71.2%), were employed (91.8%), and reported using either alcohol (86.6%) or marijuana (57.8%). The average age of the sample was 35 years, and the average level of educational attainment was the 12th grade. The sample included African American (50.9%) and Caucasian (49.1%) men averaging 7.5-year relationships. Table 1 shows the general characteristics of the sample at the pretreatment assessment.

The Batterer Intervention Program

The intervention program is cognitive–behavioral in orientation and is consistent in organization and focus to those programs described in the literature (Gelles, 2002; Holtzworth-Munroe, 2001; Rosenbaum & Leisring, 2002). The intervention program is a structured, intensive, 26-week group treatment program that focuses primarily on anger management and skills development. The intervention program incorporates three phases: (a) orientation and intake interview (two sessions), (b) psychoeducational classes (20 sessions), and (c) group therapy regarding termination (four sessions). Groups consist of approximately 15 batterers and meet 1 night each week for approximately 2 hours. This batterer treatment program incorporates confrontation, therapy, and educational components. In this setting, the common proximal events of domestic violence are directly addressed with clients.

The 20-week psychoeducational program curriculum can be broken up into three successive series of group experiences. Because most offenders share a common set of defenses (minimization, denial, and blame) that foster aggressive behavior, the first series of group sessions helps participants to recognize and overcome these defense mechanisms. In this series, which lasts 6 weeks, participants are assisted in overcoming their natural resistance to change by helping them achieve insight into their use of defense mechanisms. Thus, the first step toward modifying behavior occurs when clients recognize and accept the fact that the problem is their behavior. In the first two sessions, program rules are reviewed and reasons for using anger are explored. Participants are instructed to examine their use of anger and to identify ways to begin to change the ways that they interact with their partner and family. Importantly, the men are required to "tell their story" and explain to the other group members the reason for their arrest and referral to the intervention program. In the third and fourth sessions, the men are educated about the importance of responsibility and honesty in achieving program goals, and "roadblocks to responsibility" (i.e., minimization, denial, and blame)

are discussed. Following this educational piece, the men are required to retell their story, and group members provide corrective feedback to each other when they hear men using "roadblocks" in their retelling of the incident that led to the program referral. In the fifth and sixth sessions, the cycle of violence is explained, and the men are asked to consider their relationships from this perspective and to share their experiences with the group. At the end of this series, the concept of partnership in relationships is explored and discussed as an alternative to power and control.

The second series of sessions, which lasts 6 weeks, flows out of the fact that the belief and value systems of most batterers are very similar and foster the notion of traditional sex roles stereotypes. This series challenges the batterers' beliefs and values. The sessions are designed to help clients restructure their thinking by modifying the beliefs that promote violent behavior. In the seventh and eighth sessions, a modified form of rational emotive therapy (RET) is explained, and participants are encouraged and assisted in applying the model to their own experience. Also, in these two sessions, the concept of time-out is explained as a method to avoid the escalation that frequently accompanies arguments that result in violence. In the 9th and 10th sessions, irrational beliefs that contribute to violence (e.g., "If my partner doesn't do what I want, she is deliberately trying to make me angry") are explored and discussed, and the concept of rational self-talk is proposed as a method of avoiding the kind of irrational thinking that results in anger. In the 11th and 12th sessions, the concept of "thinking traps" (e.g., exaggeration and personalization) are presented to members, and they are required to explore in group how they have used these thinking errors to justify the violence that they have inflicted on their partner.

The final series of sessions, which last 8 weeks, is designed to help clients increase interpersonal skills by providing them with a repertoire of alternate and appropriate behaviors. In this series, skills such as problem solving, assertiveness, and negotiation are both taught and practiced in the group setting. In the 13th and 14th sessions, the concept of assertiveness is explained, and the men are required to role-play scenarios for which an assertive response is required for successful resolution of a problem situation. In the 15th and 16th sessions, "roadblocks to communication" (e.g., telling others what to do, preaching, and judging) are explained, and participants are encouraged to explore their own use of these concepts in communicating with their partners. Also, in the 14th and 15th sessions, the concept of assertiveness is continued and men are required to role-play problematic communication scenarios with each other in group. In the 17th and 18th sessions, a problem-solving model is presented and applied to the experiences of the men in group. In the 19th and 20th sessions, the concepts of trust, support, tolerance, and acceptance are presented, with an emphasis on how they relate to successful negotiation. Participants are required to discuss and role-play strategies for incorporating negotiation into their interactions with their partners.

OUTCOME MEASURES

Social Desirability

Socially desirable responding was measured using the BIDR scale (Paulhus, 1984). The BIDR is a 40-item, Likert-type scale that is equally divided between questions related to self-deception and impression management. This study used the 7-point Likert-type scale version and the continuous scoring method. Respondents who score high in self-deception are considered to believe their positive self-reports. Conversely, respondents who score high on impression management are considered to be responding in a calculated way to give a favorable impression of themselves. Self-deception and impression management scores range from 0 to 20, or low (0 to 6), moderate (7 to 13), and high (14 to 20). This two-factor model of socially desirable responding was found to be more effective than the often used attribution/denial models, which some authors suggest are equivalent measures of the same construct (Paulhus, 1984). Internal consistency reliability was tested for each subscale (Kroner & Weeks, 1996), resulting in a .84 for the impression management scale and a .73 for the self-deception scale. External validity was demonstrated through a comparison of scale differences (see Kroner & Weeks, 1996). The research suggests that the BIDR is useful in measuring socially desirable responding among offenders (Kroner & Weeks, 1996). Currently there is no information regarding the appropriateness of this measure for ethnic minorities.

Spouse-Specific Behavior

Spouse-specific behavior, both assertive and passive/aggressive, was measured using the SSAS (O'Leary & Curley, 1986). This scale, created by O'Leary and Curley (1986) was adapted from a scale designed to measure spouse-specific assertion used by Rosenbaum and O'Leary (1981). The scale measures both assertive behaviors and passive/aggressive responses directed toward their partner. Scores on the scale range from 12 to 102, with scores falling into low, moderate, and high categories. On the Assertiveness subscale, low scores range from 17 to 45, moderate from 46 to 74, and high from 75 to 102. On the

Passive/Aggressive subscale, low scores range from 12 to 31, moderate from 32 to 51, and high from 52 to 72 (O'Leary & Curley, 1986). Internal consistency reliability (Cronbach's alpha) of the SSAS was tested for each subscale (O'Leary & Curley, 1986), resulting in a .87 for the Assertiveness subscale and a .82 for the Aggressive subscale, which is considered reliable for basic research purposes (Nunnally, 1978). Currently there is no information regarding the appropriateness of this measure for ethnic minorities.

Controlling Behavior

The extent to which individuals exhibit controlling behavior with regard to a partner was measured using the CPS (Follingstad et al., 1988). The scale has seven subscales with values that range from 1 (*never or rarely*) to 4 (*quite frequently*). The seven subscales include: Jealousy Index, Verbal Criticism Index, Controlling Free Time Index, Controls Appearance or Duties Index, Physical Coercion Index, Verbal Coercion Index, and Forced to Talk About Past Sexual Behavior Index. The total control score is obtained by multiplying the mean of the subscale scores by 7, with higher scores indicating greater frequency of engaging in controlling behavior. The total control score was used in this study. The scale has demonstrated excellent internal consistency reliability (coefficient alpha = .95) and excellent criterion validity, differentiating individuals who used physical force from those who did not (Follingstad, Bradley, Laughlin, & Burke, 1999). The racial/ethnic composition of the normative sample (Follingstad et al., 1999) was roughly equivalent to the racial composition of the sample in this study.

Propensity for Abuse

The PAS was used to measure program participants' likelihood of using physical force on their partner. The scale has three factors: affective lability, recalled negative parental treatment, and trauma symptoms (Dutton, 1995). Scores range from 32 to 160, with higher scores indicating an increased likelihood that the individual will use physical force on their partner. Internal consistency reliability (Cronbach's alpha) tests of the PAS resulted in a .88, which is considered reliable for basic research purposes (Nunnally, 1978). In a study exploring the utility of the scale, Dutton (1995) found the mean score of nonviolent men to be 44.7 (SD = 11.7) and violent men to be 59.2 (SD = 17.1). Recently, Dutton, Landolt, Starzomski, and Bodnarchuk (2001) reviewed the available research on the PAS and concluded that there is ample empirical evidence to support its use with diverse groups.

RESULTS

Data Analysis

Using variables from the literature found to be effective predictors of abusive behavior, logistic regression was employed to determine which variables might have had a potential bearing on participants' treatment program completion. Logistic regression was chosen for this study over discriminant analysis for two reasons. Logistic regression does not require the assumptions of multivariate normality and equal variance–covariance matrices across groups that discriminant analysis requires, and the model effectively incorporates categorical independent variables (Hair, Anderson, Tatham, & Black, 1992; Hosmer & Lemeshow, 1989). Because 6 of the 14 predictor variables were categorical, logistic regression was considered the optimal statistical test to determine predictive efficacy of the model and to identify which of the variables in the model were significantly associated with treatment completion (Hair et al., 1992).

In the present study, this analysis strategy allowed for prediction of program completion versus dropout. In this analysis, logistic regression models the log of the odds that an individual will complete treatment as a function of a set of predictors, that are both continuous and categorical (dummy coded), and where the response is binary (completion versus noncompletion; Hosmer & Lemeshow, 1989). Using this model, it is then possible to develop a logistic regression heuristic that will allow for prediction of the outcome (see Hair et al., 1992).

Demographic Characteristics

Program completers and noncompleters were compared using a series of independent t tests and chi-square tests of significance. Analyses indicated that, relative to dropouts, program completers were older (t = 4.863, p = .000, d = .24), reported higher monthly incomes (t = 4.159, p = .000, d = .24) and longer relationships with their victim (t = 2.077, p = .038, d = .12), and they were more often referred after arrest (χ^2 = 29.222, df = 1, p = .000) and less likely to report any marijuana usage (χ^2 = 4.605, df = 1, p = .032). No significant differences were found between the two groups in level of educational attainment, employment (employed versus not employed), relationship status (married versus not married), race (Caucasian versus African American), alcohol usage (use versus no use), crack/cocaine usage (use versus no use), or prescription drug usage (use versus no use). It should be noted that conducting a series of independent t tests raises the possibility of family-wise error rate.

Psychosocial Testing Scales

Psychosocial testing scales were used to measure program completers and noncompleters on issues of socially desirable responding, assertiveness, control, and propensity for abusiveness. Table 2 presents the outcome of these comparisons. Program participants (n = 1,702) tended to score in the low to moderate range on both subscales of the BIDR: self-deception low (41.5%, n = 696), moderate (54.3%, n = 911); impression management low (42.3%, n = 705), moderate (57.7%, n = 963). More than 90% of program participants scored in the moderate (65%, n = 1,078) to high (26%, n = 431) range on the SSAS Assertiveness subscale. Conversely, only 8% (n = 137) scored in the high range on the SSAS Passive/Aggressive subscale, with the majority of participants scoring in the low (41%, n = 675) to moderate (51%, n = 847) range. In this sample, 98% (n = 1674) of program completers and noncompleters had scores on the CPS that reflected they low levels of controlling behavior. Only 2% (n = 28) of the sample received scores on the CPS that reflected that they often or frequently engaged in controlling behaviors. Program participant's scores on the PAS mirrored those of other violent men (M = 59.6, SD = 20.32), suggesting a likelihood of using physical force on their partners.

Independent t tests were also used to determine differences between treatment completers and dropouts on issues of socially desirable responding, assertiveness, control, and propensity for abusiveness. With an alpha level of .05, two-tailed independent t tests indicated that treatment completers had a greater propensity for abusiveness (t = 2.225, p = .026, d = .11) than did treatment dropouts. There were no significant differences between the two groups in terms of socially desirable responding (t = .384, p = .701), assertiveness (t = 1.703, p = .089), passive/aggressiveness (t = 1.091, p = .276), or control (t = .465, p = .642).

TABLE 2: Participant Scores on the SASS, CPS, and PAS

Scale	Completers (n = 850)	Noncompleters (n = 852)
Spouse-Specific Assertiveness Scale (SASS)		
Assertiveness		
M	63.76	65.00
SD	14.81	14.94
Missing data (*n*)	28	16
Spouse-Specific Assertiveness Scale (SASS)		
Passive/aggressive		
M	34.46	35.11
SD	12.40	12.10
Missing data (*n*)	28	15
Control of Partner Scale (CPS)		
M	78.43	78.97
SD	24.13	23.74
Missing data (*n*)	0	0
Propensity for Abusiveness Scale (PAS)*		
M	60.71	58.52
SD	20.89	19.68
Missing data (*n*)	0	0

*p < .05.

Predictive Efficacy

Reducing attrition rates for male batterer treatment programs may be possible if treatment programs were able to predict treatment completion and noncompletion and then focus program efforts to the two groups accordingly. The term *predictive efficacy* refers to the agreement between expected and observed values. The predictive efficacy of a measure is maximized when the expected and observed values are approximate (Visher, Lattimore, & Linster, 1991). The observed instances of program completion were 49.6% (n = 822) for this sample (n = 1656, 97.3% of sample) as compared to the percentage of males predicted to complete treatment (47.8%, n = 792). The percentage of correctly predicted treatment completers (28.9%, n = 478) and correctly predicted treatment dropouts (31.4%, n = 520) resulted in correct classification in 60.3% of the sample. The model incorrectly classified treatment completion in 39.7% (n = 658) of the cases, with 18.9% (n = 314) false positives and 20.8% (n = 344) false negatives. Equally important is the extent to which the model improves, over chance, the identification of treatment completers. To assess the degree to which the measure improved over the chance level of predictive efficacy, the Relative Improvement Over Chance Index (RIOC) was used (see Loeber & Dishion, 1983). Analysis indicated that the model resulted in a 10.7% improvement over chance (49.6% versus 60.3%) in the prediction of treatment completion for the sample.

This study employed an analysis strategy that allowed for simultaneous entry of the independent variables. The estimated coefficients of the logistic regression model are presented in Table 3. The sign of the logistic coefficients (positive or negative) indicate increases or decreases in the dependent variable, with a one-unit change in each independent variable (Hair et al., 1992). For example, the regression coefficient for age was significant, indicating that age of the batterer affected the likelihood of treatment completion when the control variables were held constant. For each unit increase in age, the odds of the participant completing treatment are increased by a factor of 1.021, when all other independent variables are held constant. It should be noted that the sample for the logistic regression was 1,656 (97.3%), reduced by 46 cases from the total sample due to missing data on the independent variables. The model, however, satisfies the ratio of 5 cases per variable suggested by Hair et al. (1992).

The overall fit of the model was significant (χ^2 = 86.965, df = 14, p = .000). When treatment completion was modeled

TABLE 3: Logistic Regression Analysis of Treatment Completers and Noncompleters—26-Week Treatment Program

Variable	Logistic Coefficients	*SE*	Wald χ^2	Exp(*B*)
Demographic factors				
Age	.021***	.006	11.737	1.021
Race	.054	.111	0.232	1.055
Marital status	–.073	.108	0.448	0.930
Relationship length	.000	.001	0.089	0.765
Level of education	–.003	.024	0.012	0.912
Average income	.000***	.000	15.067	1.000
Alcohol/drug usage factors				
Alcohol	.061	.164	0.138	10.63
Marijuana	.072	.118	0.370	1.074
Crack/cocaine	.306**	.142	4.605	1.358
Psychological factors				
Spouse-specific assertiveness	–.005	.004	1.647	0.995
Spouse-specific passive/aggressiveness	–.006	.005	1.252	0.994
Control of partner	–.001	.003	0.050	0.999
Propensity for abusiveness	.011	.003	12.617	1.011
After arrest	.886***	.185	23.069	0.412
Constant	–1.414			
Log likelihood	2208.65			
Model χ^2	86.965[a]***			

NOTES: $n = 1{,}656$.
a. $df = 14$.
*$p < .05$. **$p < .01$. ***$p < .001$.

TABLE 4: Logistic Regression Analysis of Treatment Completers and Noncompleters—16-Week Treatment Program

Variable	Logistic Coefficients	*SE*	Wald χ^2	Exp(*B*)
Demographic factors				
Age	0.082**	.027	9.518	1.085
Alcohol/drug usage factors				
Alcohol	–0.595	.704	0.714	0.234
Marijuana	0.115	.557	0.043	1.122
Crack/cocaine	0.732	.673	1.183	2.079
Prescription drugs	0.470	.489	0.925	1.600
Psychological factors				
Spouse-specific assertiveness	–0.038*	.018	4.459	0.962
Spouse-specific passive/aggressiveness	–0.030	.027	1.293	0.970
Control of partner	0.022	.013	2.637	1.022
Propensity for abusiveness	–0.014	.014	1.037	0.986
Referred after arrest	1.453*	.723	4.037	0.234
Currently employed	0.449	.810	0.307	1.567
Constant	–0.230			
Log likelihood	124.443			
Model χ^2	31.827[a]***			

NOTES: $n = 115$.
a. $df = 11$.
*$p < .05$. **$p < .01$. ***$p < .001$.

to be dependent on the 14 factors of the model, five of the predictive variables were significant. The treatment completer's age (Wald $\chi^2 = 11.737$, $df = 1$, $p = .001$), participant score on the PAS (Wald $\chi^2 = 12.617$, $df = 1$, $p = .000$), average income (Wald $\chi^2 = 15.067$, $df = 1$, $p = .000$), crack/cocaine usage (Wald $\chi^2 = 4.605$, $df = 1$, $p = .032$), and whether the participant was referred after arrest (Wald $\chi^2 = 23.069$, $df = 1$, $p = .000$) significantly contributed to the prediction of treatment completion among male batterers.

Testing the 16-Week Model

In a similar study conducted on the 16-week treatment program model, Buttell and Carney (2002) created a predictive model that resulted in correct classification in 74.8% of the sample, (47.0%, $n = 54$ predicted treatment completers; 27.8%, $n = 32$ predicted treatment dropouts), and a 33.1% improvement over chance (41.7% versus 74.8%) in the prediction of treatment completion for the sample.

In addition, the overall fit of the model was significant ($\chi^2 = 31.827$, $df = 11$, $p = .001$). When treatment completion was modeled to be dependent on the 11 factors of the model, three of the predictive variables were significant (see Table 4). The treatment completer's age (Wald $\chi^2 = 9.518$, $df = 1$, $p = .002$), participant score on the SSAS Assertiveness subscale (Wald $\chi^2 = 4.459$, $df = 1$, $p = .035$), and whether the participant was referred after arrest (Wald $\chi^2 = 4.037$, $df = 1$, $p = .045$) significantly contributed to the prediction of treatment completion among male batterers.

TABLE 5: Logistic Regression Analysis of Treatment Completers and Noncompleters 26-Week Treatment Program Using 16-Week Regression Model

Variable	Logistic Coefficients	*SE*	Wald χ^2	Exp(*B*)
Demographic factors				
Age	.026***	.006	20.791	1.027
Alcohol/drug usage factors				
Alcohol	.112	.167	0.449	1.119
Marijuana	.097	.118	0.685	1.102
Crack/cocaine	.260	.141	3.387	1.297
Prescription drugs	–.014	.104	0.018	0.986
Psychological factors				
Spouse-specific assertiveness	–.005	.004	2.051	0.995
Spouse-specific passive/aggressiveness	–.004	.005	0.572	0.996
Control of partner	–.001	.003	0.138	0.999
Propensity for abusiveness	.011***	.003	12.910	1.011
Referred after arrest	.835***	.183	20.812	0.434
Currently employed	–.286	.188	2.309	0.751
Constant	–1.158			
Log likelihood	2201.63			
Model χ^2	69.102a***			

NOTES: $n = 1{,}638$.
a. $df = 14$.
*$p < .05$. **$p < .01$. ***$p < .001$.

When applied to the 26-week treatment program sample, the predictive model created for the 16-week treatment program produced a predictive model for the 26-week group of completers and dropouts that resulted in correct classification in 57.3% of the sample (28.6%, $n = 469$ predicted treatment completers; 28.7%, $n = 470$ predicted treatment dropouts). Also, the applied model resulted in a 7.5% improvement over chance (49.8% versus 57.3%) in the prediction of treatment completion for the 26-week sample as compared to the predictive model that resulted in a 33.1% improvement over chance (41.7% versus 74.8%) in the prediction of treatment completion for the 16-week sample.

However, the overall fit of the applied model was significant ($\chi^2 = 69.102$, $df = 11$, $p = .000$). When treatment completion was modeled to be dependent on the 11 factors of the applied model, three of the predictive variables were significant (see Table 5)—two of which were also significant with the 16-week group. The treatment completer's age (Wald $\chi^2 = 20.791$, $df = 1$, $p = .000$) and whether the participant was referred after arrest (Wald $\chi^2 = 20.812$, $df = 1$, $p = .000$) significantly contributed to the prediction of treatment completion among male batterers as well as the participant score on the PAS (Wald $\chi^2 = 12.910$, $df = 1$, $p = .000$)

DISCUSSION AND APPLICATIONS TO PRACTICE

The results of this study are important because they represent the first investigation of the usefulness of BIP attrition models developed on programs of shorter durations with the programs of longer duration mandated by state standards. Not surprisingly, given the confusing nature of the batterer-intervention attrition literature discussed previously, the findings of this study failed to identify many differences between treatment completers and dropouts on either demographic or psychological variables. Although the analysis suggested that, relative to dropouts, treatment completers (a) were older, (b) reported higher monthly incomes, (c) were involved in longer relationships with their victims, (d) were more often referred after arrest, (e) were less likely to report marijuana usage, and (f) demonstrated a greater propensity for abusiveness, other demographic and psychosocial variables did not distinguish between the two groups. Unfortunately, the findings from this study failed to help clarify the confusing relationship between demographic variables and program attrition present in the national literature. Specifically, these findings are inconsistent with some previous research in the field suggesting that educational attainment, employment status, self-reported alcohol use, and marital status are important predictors of treatment attrition (Buttell & Carney, 2002; Cadsky et al., 1996; Chang & Saunders, 2002; Daly et al., 2001; DeMaris, 1989; Gruznski & Carrillo, 1988; Hamberger & Hastings, 1989; McCloskey et al., 2003; Scott, 2004). However, the findings are consistent with other investigations into batterer-treatment-program attrition that found little or no relationship between demographic variables and premature dropout (Buttell & Pike, 2002; Chen, Bersani, Myers, & Denton, 1989; DeHart, Kennerly, Burke, & Follingstad, 1999; DeMaris, 1989; Gruznski & Carrillo, 1988; Hamberger & Hastings, 1986, 1991; Hamberger et al., 2000; Jones & Gondolf, 2002).

Taken as a whole, these findings would seem to provide strong support for Buttell and Carney's (2004) assertion that individual BIPs need to consider problems of program attrition from their unique situation and location. Interestingly, what is most encouraging about these findings is that they suggest the possibility that the most violent batterers are being retained in this treatment program at a greater rate than less violent batterers. Specifically, if domestic violence is viewed along a continuum from emotional abuse on one end to rape on the other end, then these findings suggest that the men more likely to complete treatment are those who self-report a greater propensity for being physically abusive toward their female partners. In either case, the findings from the logistic regression model are encouraging because they suggest ways in which dropouts can be identified at the pretreatment assessment, with an eye toward enhancing retention.

Perhaps the most important findings from this study relate to the logistic regression model and the replication of the 16-week logistic regression model with the 26-week program data. Specifically, the findings demonstrate that a batterer intervention program can develop a reliable method for identifying men at greatest risk of dropping out of their specific program. This finding is important because it suggests that there is no "one-size-fits-all" model to approaching batterer intervention attrition in terms of specific demographic or psychological variables. In the past, the confusing nature of the batterer intervention attrition literature prevented meaningful synthesis and analysis. Consequently, individual intervention programs were left with very little good information about how they could increase retention given each program's unique situation in regard to factors like geographic location and judicial support. These findings suggest that individual programs can partner with social work researchers to discover ways to make use of existing data in both identifying and preventing dropouts. In this regard, perhaps the most hopeful finding from this study was that it was able to create a significant regression model, using a different set of instruments than those used by other researchers (e.g., Buttell & Pike, 2002). This suggests that individual programs can identify local predictors of attrition and that this identification is not dependent on a particular set of instruments. Equally important is the fact that the model captured 97% of the sample while correctly classifying 60% of the sample. Although the model only represents an 11% improvement over chance, from the standpoint of clinical significance, it is still quite important. Specifically, if the BIP is able to use the model to retain even one additional man that would have otherwise dropped out, it would be difficult to argue to the partners and children of these men that their retention in treatment was unimportant.

In regard to the replication of the 16-week model with the 26-week program data, two issues are important. First, given the fact that most BIPs nationally do not have staff researchers or access to powerful statistical software, the benefit of the discovery that a significant model was created using different instruments from those used in other research is that it might have wide applicability nationally. To illustrate, consider the application of these findings for the intervention program that was the site for this study. The agency has been using the data in this study to create a pretreatment assessment of the men beginning the intervention program and for nothing else. Prior to the collaborative relationship with the researchers, the intervention program did not have the staff resources or training to use that data to create a logistic regression model to predict dropout. However, the agency did have Microsoft Excel and the authors have created a scoring strategy, based on the output of the logistic regression model, and trained agency staff on how to use the data so that they can make use of the regression model without having to purchase any new software or learn any new statistical methods. Having a predictive model in place makes this type of prospective identification possible, and if men at risk of dropping out are identified early, any number of measures to enhance their retention could be tried. For example, the treatment program that was the site for this study has used the predictive model to identify men at risk of dropping out and then provide them with some crisis-oriented services. The agency has developed a plan to attempt to retain these men in treatment by having a staff member perform individual, crisis-oriented therapy with these clients to explore with them any issues that may be contributing to their risk of dropping out of the program. In this regard, this study might serve as a model for batterer intervention programs nationally in their efforts to identify local norms and predictors for attrition. However, the advent of a new state law mandating 26 weeks of treatment for BIPs called into question the utility of the model that was developed on the 16-week program data.

The most interesting outcome of this study was the discovery that some of the same variables used to predict attrition in the 16-week program were also useful in predicting attrition in the 26-week program. This issue is increasingly important because BIPs are legislatively mandated to increase program length—in this case, by 38%. Consequently, as discussed previously, it was not clear whether the models developed on programs of shorter duration would have any utility in predicting attrition involving longer programs because judicial

support for the longer program may wane or interact with some unknown variable and affect the models, predictive efficacy. The results of this study suggest that the tools for creating successful models with programs of longer duration might be connected to the models that were developed on shorter programs. This is of critical importance for the reason just discussed: namely, that most BIPs do not have the expertise or staff resources to create their own models. These findings suggest that what may be required for the successful identification of men at risk of prematurely dropping out of the longer BIP are some of the same variables that BIPs are already collecting. Consequently, social work researchers can work collaboratively with BIPs, in the ways just described, to make use of existing data collections mechanisms and software to reliably identify men at risk of premature dropout.

Limitations

There are two drawbacks to this study that limit the conclusions that can be drawn from it, and it is important to keep these limitations in mind when evaluating the findings. First, this study employed a sample of batterers drawn from a predominately rural, Southern state. It is clear that these batterers are not representative of batterers in general, and the results of this study may not be applicable to batterers in different geographic regions and clinical settings. Second, as discussed previously, the predictive model is only useful for the men attending this particular treatment program. Other treatment programs around the country will have to develop their own models based on the characteristics of the men attending their treatment programs.

REFERENCES

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

Batterer Intervention Services Coalition Michigan. (2006). *State standards*. Retrieved September 1, 2006, from www.biscmi.org\other_resources\state_standards.html

Buttell, F., & Carney, M. (2002). Predictors of attrition among batterers court-ordered into treatment. *Social Work Research*, *26*, 31-34.

Buttell, F., & Carney, M. (2004). A multidimensional assessment of a batterer treatment program: An alert to a problem? *Research on Social Work Practice*, *14*, 93-101.

Buttell, F., & Pike, C. (2002). Investigating predictors of treatment attrition among court-ordered batterers. *Journal of Social Service Research*, *29*, 53-68.

Cadsky, O., Hanson, K., Crawford, M., & Lalonde, C. (1996). Attrition from a male batterer treatment program: Client-treatment congruence and lifestyle instability. *Violence and Victims*, *11*, 51-64.

Chang, H., & Saunders, D. (2002). Predictors of attrition in two types of group programs for men who batter. *Journal of Family Violence*, *17*, 273-292.

Chen, H., Bersani, C., Myers, S., & Denton, R. (1989). Evaluating the effectiveness of a court-sponsored abuser abatement program. *Journal of Family Violence*, *4*, 309-322.

Daly, J., & Pelowski, S**.** (2000). Dropout from groups for men who batter: A review of studies with implications for research and practice. *Violence and Victims*, *15*(2), 137-160.

Daly, J., Power, T., & Gondolf, E. (2001). Predictors of batterer program attendance. *Journal of Interpersonal Violence*, *16*, 971-991.

Davis, R., & Taylor, B. (1999). Does batterer treatment reduce violence? A synthesis of the literature. *Women and Criminal Justice*, *10*, 69-93.

Davis, R., Taylor, B., & Maxwell, C. (1998). *Does batterer treatment reduce violence? A randomized experiment in Brooklyn* (NIJ Final Report, Grant No. 94-IJ-CX-0047). Washington, DC: National Institute of Justice (NIJ)/National Criminal Justice Reference Service.

DeHart, D., Kennerly, R., Burke, L., & Follingstad, D. (1999). Predictors of attrition in a treatment program for battering men. *Journal of Family Violence*, *14*, 19-35.

DeMaris, A. (1989). Attrition in batterers' counseling: The role of social and demographic factors. *Social Service Review*, *63*, 142-154.

Dunford, F. (2000). The San Diego Navy Experiment: An assessment of interventions for men who assault their wives. *Journal of Consulting and Clinical Psychology*, *68*, 468-476.

Dutton, D. (1995). A scale for measuring propensity for abusiveness. *Journal of Family Violence*, *10*, 203-221.

Dutton, D. G., Landolt, M., Starzomski, A., & Bodnarchuk, A. (2001). Validation of the Propensity for Abusiveness Scale in diverse male populations. *Journal of Family Violence*, *16*, 69-73.

Feder, L., & Forde, D. (2000). *A test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward experiment* (NIJ Final Report). Washington, DC: NIJ.

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

Follingstad, D. R., Bradley, R. G., Laughlin, J. E., & Burke, L. (1999). Risk factors and correlates of dating violence: The relevance of examining frequency and severity levels in a college sample. *Violence and Victims*, *14*, 365-380.

Follingstad, D. R., Rutledge, L., McNeill-Harkins, K., & Polek, D. (1988). Factors related to physical violence in dating relationships. *Victimology International Journal*, *3*, 169-182.

Gelles, R. (2002). Standards for programs for men who batter? Not yet. *Journal of Aggression, Maltreatment & Trauma*, *5*, 11-20.

Gondolf, E. (2001). Limitations of experimental evaluations of batterer programs. *Trauma, Violence, and Abuse, 2,* 79-88.

Gruznski, R., & Carrillo, R. (1988). Who completes batterer treatment programs? An empirical investigation. *Journal of Family Violence*, *3*, 141-150.

Hair, J. F., Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (1992). *Multivariate data analysis with readings*. New York: Macmillan.

Hamberger, L. K., & Hastings, J. E. (1986, November). *Characteristics of male spouse abusers: Is psychopathology part of the picture?* Paper presented at American Society of Criminology, Atlanta, GA.

Hamberger, K., & Hastings, J. (1988). Skills training for treatment of spouse abusers: An outcome study. *Journal of Family Violence*, *3*, 121-130.

Hamberger, K., & Hastings, J. (1989). Counseling male spouse abusers: Characteristics of treatment completers and drop-outs. *Violence and Victims*, *4*, 275-286.

Hamberger, K., & Hastings, J. (1991). Personality correlates of men who batter and nonviolent men: Some continuities and discontinuities. *Journal of Family Violence*, *6*, 131-147.

Hamberger, K., Lohr, J., & Gottlieb, M. (2000). Predictors of treatment dropout from a spouse abuse abatement program. *Behavior Modification*, *24*, 528-552.

Holtzworth-Munroe, A. (2001). Standards for batterer treatment programs: How can research inform our decisions? *Journal of Aggression, Maltreatment & Trauma*, *5*, 165-180.

Hosmer, D. E., & Lemeshow, S. (1989). *Applied logistic regression*. New York: John Wiley.

Jones, A., & Gondolf, E. (2002). Assessing the effect of batterer program completion of reassault: An instrumental variable analysis. *Journal of Quantitative Criminology*, *18*, 71-98.

Kroner, D. G., & Weeks, J. R. (1996). Balanced inventory of desirable responding: Factor structure, reliability, and validity with an offender sample. *Personality and Individual Differences*, *21*, 323-333.

Kyriacou, D., Anglin, D., Taliaferro, E., Stone, S., Tubb, T., Linden, J., et al. (1999). Risk factors for injury to women from domestic violence against women. *New England Journal of Medicine*, *341*, 1892-1898.

Levesque, D. (1999). *Violence desistance among battering men: Existing interventions and the application of the transtheoretical model of change*. Unpublished doctoral dissertation, University of Rhode Island, Kingston.

Loeber, R., & Dishion, T. J. (1983). Early predictors of delinquency: A review. *Psychological Bulletin*, *94*, 68-99.

Maiuro, R., Hagar, T., Lin, H., & Olson, N. (2001). Are current state standards for domestic violence perpetrator treatment adequately informed by research? A question of questions. *Journal of Aggression, Maltreatment & Trauma*, *5*, 21-44.

McCloskey, K., Sitaker, M., Grigsby, N., & Malloy, K. (2003). Characteristics of male batterers in treatment: An example of a localized program evaluation concerning attrition. *Journal of Aggression, Maltreatment & Trauma*, *8*, 67-95.

Morrison, S., & Nesius, A. (2003, September). *Findings from an evidenced-based review of batterer prevention and intervention programs*. Paper presented at the Eighth International Conference on Family Violence, San Diego, CA.

Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.

O'Leary, K. D., & Curley, A. D. (1986). Assertion and family violence: Correlates of spouse abuse. *Journal of Marital and Family Therapy*, *12*, 281-289.

Paulhus, D. L. (1984). Two component models of social desirability responding. *Journal of Personality and Social Psychology*, *46*, 598-609.

Rennison, C., & Welchans, S. (2000). *Intimate partner violence* (Bureau of Justice Statistics Report, NCJ 178247). Washington, DC: U.S. Department of Justice.

Rosenbaum, A., & Leisring, P. (2002). A group intervention programs for batterers. *Journal of Aggression, Maltreatment & Trauma*, *5*, 57-71.

Rosenbaum, A., & O'Leary, K. D. (1981). Marital violence: Characteristics of abusive couples. *Journal of Consulting and Clinical Psychology*, *49*, 63-71.

Scott, K. (2004). Stage of change as a predictor of attrition among men in a batterer intervention program. *Journal of Family Violence*, *19*, 37-47.

Tjaden, P., & Thoennes, N. (2000). *Extent, nature, and consequences of intimate partner violence: Findings from the National Violence Against Women Survey* (Publication No. NCJ 181867). Washington, DC: U.S. Department of Justice.

Visher, C. A., Lattimore, P. K., & Linster, R. L. (1991). Predicting the recidivism of serious youthful offenders using survival models. *Criminology*, *29*, 329-362.