

Investigating the Differential Effectiveness of a Batterer Treatment Program on Outcomes for African American and Caucasian Batterers

Frederick P. Buttell

Tulane University

Cathy K. Pike

Indiana University—

Purdue University, Indianapolis

Objective: This study (a) evaluated a batterer intervention program (BIP) by investigating changes in psychological variables (i.e., truthfulness, violence, lethality, control, alcohol use, drug use, and stress-coping abilities) between pretreatment and posttreatment assessments in a sample of court-mandated batterers and (b) investigated the differential effectiveness of this same BIP for African American and Caucasian batterers. Method: The study employed a one-group pretest/posttest design, with 12-month follow-up data, to investigate changes in Domestic Violence Inventory (DVI) scores among 91 men, 57% African American, court ordered into a BIP. Results: Analysis indicated that (a) court-ordered batterers demonstrate significant changes, in the desired direction, on psychological variables related to domestic violence, as a result of participation in a court-mandated BIP and (b) there was no significant difference in changes on these psychological variables between African American and Caucasian batterers. Conclusion: Implications of the findings for enhancing intervention efforts with court-ordered batterers were discussed.

Keywords: *court-ordered batterers*

Despite more than 20 years of intervention efforts with battered women, domestic violence remains the most common cause of nonfatal injury to women in the United States (Kyriacou et al., 1999). Although the most recent statistics compiled by the Bureau of Justice Statistics suggest that the rate of partner violence directed toward women dropped 21% between 1993 and 1998, these same statistics indicate that there were still more than 875,000 violent incidents directed at women by their partners in 1998 (Rennison &

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

Research on Social Work Practice, Vol. 13 No. 6, November 2003 675-692
DOI: 10.1177/1049731503254055
© 2003 Sage Publications

Welchans, 2000). One of the explanations for the decline in the rate of partner violence is the proliferation of batterer intervention programs (BIPs) across the country (Gondolf, 1997). The premise of BIPs is that to be successful, efforts to eliminate family violence must focus on the individual who has the problem. In instances involving domestic violence, this is the person who initiates the assaultive behaviors in the family (i.e., the batterer). The reduction in violent incidents between 1993 and 1998 would suggest that these programs are having the desired effect. However, the large number of women who continue to be assaulted annually also suggests that there remains considerable room for improvement.

In the effort to improve intervention efforts for batterers, a growing body of literature has focused on characteristics of men who batter and on the intervention programs directed at changing their violent behaviors. For the purposes of this discussion, the terms *battering* and *domestic violence* are used interchangeably to refer to a multidimensional construct that includes physical violence, emotional abuse, and controlling behaviors. Relative to the current discussion, reviews of the BIP literature have consistently indicated that (a) 60% to 80% of male batterers who complete treatment are no longer physically abusive toward their partners at the conclusion of the treatment program (Eisikovits & Edleson, 1989; Gondolf, 1997; Holtzworth-Munroe, Bates, Smultzer, & Sandin, 1997; Rosenfeld, 1992; Tolman & Bennett, 1990) and (b) treatment programs ignore cultural differences among batterers and treat them all with the same cognitive-behavioral approach (Bennett & Williams, 2001; Williams, 1992, 1994; Williams & Becker, 1994). Although the first finding appears encouraging, there are numerous methodological limitations that limit the confidence that can be placed in these positive, single-site program evaluations (for an excellent discussion, see Rosenfeld, 1992). For example, many evaluations employ self-report data in their analysis of treatment effect. Specifically, many programs simply ask treatment completers if they are currently engaging in any violent or abusive behaviors at the termination of treatment services. Unfortunately, the correct answers to these questions are obvious and, even if participants wanted to respond honestly, they may fear negative consequences from reporting that they are still engaging in violent or controlling behaviors (e.g., a recommendation of continued treatment). Other serious limitations include lack of posttreatment follow-up data (e.g., rearrest data or victim interviews) to evaluate if any changes observed at the posttest evaluation are maintained over time and the fact that very few evaluations attempt to link aspects of the intervention program with the data collected at the posttest evaluation.

In addition to these methodological limitations, a more serious omission in the BIP literature is the absence of empirical studies investigating the

differential outcome of these treatment programs for African American and Caucasian batterers (Bennett & Williams, 2001). This omission is troubling because there is clear evidence that domestic violence is occurring at a high rate in African American relationships (Oliver, 2000; Plass, 1993; Wyatt, Axelrod, Chin, Carmona, & Loeb, 2000) and many African American batterers drop out of treatment prematurely (Gondolf, 1997; Williams, 1995). Despite the need for culturally sensitive services for this population, research has consistently demonstrated the absence of such interventions in the field and no studies have investigated the possible differential impact of the standardized cognitive-behavioral treatment program on African American and Caucasian batterers (Bennett & Williams, 2001; Williams, 1992, 1994; Williams & Becker, 1994). Unfortunately, it is unclear whether the standardized treatment program works better, worse, or equally well for African American batterers. Recently, this question has taken on added importance, as there is a national legislative trend to institutionalize the "one size fits all" model (Moore, Greenfield, Wilson, & Kok, 1997; Williams, 1992). Consequently, it is imperative that treatment programs begin to evaluate whether this one size fits all model works for diverse client populations.

Given the limitations present in the BIP evaluation literature described here, the purpose of the present study was to (a) evaluate a BIP by investigating changes in psychological variables (i.e., truthfulness, violence, lethality, control, alcohol use, drug use, and stress-coping abilities) between pretreatment and posttreatment assessments in a sample of men court mandated into treatment and (b) to investigate the differential effectiveness of this same BIP for African American and Caucasian batterers. Additionally, this study evaluated arrest records for a period of 12 months following treatment completion to determine the association between changes on these psychological variables and recidivism. In brief, this study investigated two research questions: (a) Does the current standardized treatment program for batterers significantly alter psychological variables related to domestic violence and do changes on these variables predict rearrest at 12 months posttreatment? and (b) Does the current standardized treatment program for batterers differentially affect African American and Caucasian batterers?

METHOD

Participants

A total of 100 men arrested in Tuscaloosa county and ordered by the court into a BIP (treatment program) under the Alabama Law Enforcement

Protection Act were recruited to participate in this study. All potential participants received an explanation of the nature of the research and a written request for participation in the study. Those individuals who agreed to participate in the study signed a copy of the informed consent document, which described both the purpose of the study and any anticipated risks/benefits of participation. Nine of the individuals recruited for the study refused to participate. Therefore, the participants in this study were 91 men beginning their treatment for domestic violence offenses.

Data Collection

Data collection involved gathering Domestic Violence Inventory (DVI) (Risk & Needs Assessment, Inc., 1997a, 1997b) scores from 91 men arrested in Tuscaloosa County and ordered into treatment by the court. Participants were tested during their initial assessment interview for the program and on the last night of group treatment with the DVI, and they were also asked to complete a brief demographic information questionnaire. Individuals participated on two occasions for approximately 45 minutes each time. To avoid any misrepresentation of program effectiveness, rearrest records for program participants were checked at 12 months following treatment completion by the court advocate from Turning Point, the local shelter for battered women.

The BIP

The site for this study was a nonprofit organization that has been providing counseling services to the Tuscaloosa, Alabama, community for the past 25 years. Since 1989, this agency has been treating convicted domestic violence offenders. The program is part of a collaborative effort (involving the police, the court, and agency staff members) aimed at intervening in violent relationships to reduce the occurrence of domestic violence in the community. The creation of the BIP was a response to the passage of the Alabama Law Enforcement Protection Act of 1989. The Law Enforcement Protection Act, commonly referred to as the "warrantless arrest act," allows police who respond to a domestic violence call to arrest the abuser and press charges themselves. In these cases, the victim does not have to file a warrant against the abuser before an arrest is made. This law removes from the victim the burden of pressing charges and has resulted in a substantial increase in the number of domestic violence arrests and convictions in Alabama. The intervention program was created to provide a treatment alternative to incarceration as a sentencing option for judges. Due to the passing of the Law Enforcement Protection Act, the police have engaged in a proarrest policy, the courts have

mandated counseling for the batterer as part of the criminal sentence, and the agency staff members have provided the treatment component.

The intervention program is cognitive-behavioral in orientation and is consistent in organization and focus to those programs described in the literature (Williams, 1992; Gondolf, 1997). The intervention program is a structured, intensive, 12-week, group treatment program that focuses primarily on anger management and skills development (Deavers, 1998). Groups consist of approximately 20 batterers and are co-led by two group leaders in a male/female, Black/White team. Groups meet one 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 and they are given an opportunity to make changes that will positively affect their personal relationships with others.

The program curriculum can be broken up into three successive series of group experiences. Because most batterers initially refuse to accept responsibility for their violent behavior, the first series of group sessions helps participants become accountable by accepting responsibility for their violence and recognizing how their violence negatively affects both their lives and the lives of their families. In this series, participants are assisted in overcoming their natural reluctance to change by helping them understand how their violence has undermined healthy relationships and interfered with the type of relationship they desire with their partners and families. Thus, the first step toward modifying behavior occurs when clients recognize and accept the fact that the problem is their behavior. The second series of sessions flows out of the fact that the belief and value systems of most batterers are very similar and foster the notion of traditional sex role 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. The final series of sessions 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. Typically, the first series lasts 3 weeks and the second and third series are approximately 3 weeks and 6 weeks in length, respectively.

The DVI

The pretest-posttest version of the DVI was used to measure a number of factors related to domestic violence (Risk & Needs Assessment, Inc., 1997a). This inventory was initially developed in 1990. The 142-item pretest-posttest

version contains 6 scales: Truthfulness, Violence (lethality), Control, Alcohol, Drug, and Stress Coping Abilities. The Truthfulness scale measures the extent to which respondents' scores are consistent across the inventory. Respondents with Truthfulness scores at or above the 90th percentile represent inaccuracy, possibly caused by attempted "faking," denial, misreading of questions, or other reading difficulties. The Violence scale measures the extent to which respondents represent a danger to themselves or others. The Control scale measures emotionally controlling and other abusive behaviors. The Alcohol and Drug scales measure use and abuse of alcohol and drugs, problems that often are associated with domestic violence. The final scale, Stress Coping Abilities, measures subjective stress in relation to a respondent's ability to cope with stress.

The scale items use a mix of mostly "true-false" along with some Likert-type response categories. Additionally, some background variables related to domestic violence are included in the DVI scoring methods (e.g., alcohol, drug, and domestic violence arrests). The developers maintain an ongoing database that contains data from domestic violence treatment programs that use their inventory. Each scale is scored in terms of the range of risk for domestic violence, with the following categories and levels of risk: 0-39th percentile, low risk; 40th-69th percentile, medium risk; 70th-89th percentile, problem risk; and, 90th-100th percentile, severe problem risk.

A number of psychometric studies, ranging from 1990 to the present, have evaluated the reliability and validity of the DVI. Respondents in the psychometric studies have included college students (to test the Truthfulness scale), domestic violence defendants (males and females), those who have been convicted or adjudicated as domestic violence offenders, and incarcerated male prison inmates. Internal consistency reliability coefficients (Cronbach's alpha) have ranged from .83 to .95 across the scales throughout their development. The DVI has been evaluated over time with regard to its known groups, concurrent criterion, convergent construct, and discriminant construct validity. Over time, the DVI has demonstrated satisfactory to excellent levels of validity. Two additional studies examined gender differences. Gender-specific norms subsequently were developed, based on the findings that females tend to score lower than males on the Alcohol, Drug, and Violence scales. Consequently, the norms for women reflect these gender differences and are incorporated in the scoring program (Risk & Needs Assessment, Inc., 1997b).

Although the DVI relies on self-report data, the DVI represents a significant improvement over the types of self-report data utilized in other studies. For example, many other evaluations have simply asked treatment completers if they are still abusive at the conclusion of the treatment program.

Conversely, the DVI is a multi-item instrument that investigates social desirability, adopts a multidimensional construct of abuse that incorporates physical violence, emotional abuse, and controlling behaviors, and assesses for alcohol and drug abuse. In this regard, the DVI is far superior to other forms of self-report data.

RESULTS

Missing Data

A few variables contained some missing data for the total sample: educational level, age at first conviction, number of alcohol arrests, and number of drug arrests. Of these, the number of times arrested and educational level had one and two instances of missing data, respectively. Age at first conviction, number of drug arrests, and number of alcohol arrests each contained five instances of missing data for the total sample. There were no instances in the total sample of missing data for age, number of times arrested for domestic violence violations, or for any of the DVI scale scores. For all statistical analyses, cases with missing data were deleted listwise.

Sample Characteristics

Ninety-one males participated in the domestic violence program. Of these, 8 clients had pretest truthfulness scale scores that exceeded the recommended score for analysis (90th percentile and above) and were excluded from statistical tests. However, they were included in the computations of descriptive statistics on sociodemographic data.

Clients' ages ranged from 19 to 59 with a typical age of almost 31 (median = 29). Fifty-seven percent ($n = 52$) of the program's clients were African American. Comparatively more clients reported that they were married ($n = 36$, 40.9%). Nearly 88% of clients reported that they were either married, living together, or dating. The median annual income was between \$10,001 and \$20,000. Most clients had completed high school. Years of education ranged from 6 to 17 with an average of 11.85 and a median of 12. Table 1 provides additional information about the respondents' demographic characteristics.

Clients reported on a number of factors that are related to domestic violence: the number of times arrested, the number of times arrested for alcohol and drugs, the number of times arrested for domestic violence, and age at first conviction. The number of times clients reported they had been arrested ranged from none to 23, with the typical number being 3.44 (median = 3).

TABLE 1: Frequency Distributions of Demographic Variables

<i>Characteristic</i>	<i>n</i>	<i>%</i>
Age ^a		
19-28 years	41	45.1
29-39 years	37	40.7
40-59 years	13	14.3
Ethnicity		
African American	52	57.8
Caucasian/White	38	42.2
Missing	1	1.1
Marital Status		
Married	36	40.9
Living with Other	25	28.4
Dating	16	17.6
Other	11	12.5
Missing	3	3.3
Education Level		
Less than high school	32	36.0
12 years	34	38.2
13-17 years	23	25.3
Missing	2	2.2
Annual Income		
0-10,000	23	31.9
10,000-20,000	26	36.1
20,001-30,000	16	22.2
30,001-40,000	6	8.3
40,001-50,000	1	1.4

a. Mean = 30.65, Median = 29.0, *SD* = 8.37.

The number of alcohol arrests on average was almost 1.0 (median = 0), with a range of 0 to 7. Clients of this program were less likely to have been arrested for drugs than alcohol. This variable ranged from 0 to 2 and had a median of 0. As expected, the most frequently reported reason for arrest was for domestic violence, with a range from 0 to 23 and an average of 1.46 (median = 1.0). The typical age at first conviction was 23.65, with a median of 21 and a range from 11 to 54. Table 2 provides additional information on factors related to domestic violence and descriptive statistics by race and for the total sample for the DVI scales.

Information with regard to treatment completion and rearrest within the year was available for 90 of the 91 clients. Sixty-six clients had completed the program and 25 had not, yielding a completion rate of 72.5%. Seventy-two (80%) of the clients for whom this data was available had not been rearrested

TABLE 2: Descriptive Statistics for Domestic Violence-Related and DVI Variables

<i>Variable</i>	<i>Mean</i>	<i>Median</i>	<i>Standard Deviation</i>
Number of arrests	3.44	3.00	3.28
Age at first conviction	23.65	21.00	8.26
Number of domestic violence arrests	1.47	1.00	2.43
Number of alcohol arrests	4.13	0.00	17.87
Number of drug arrests	3.58	0.00	17.93
Pretest truthfulness	52.20	55.00	22.31
Caucasian	52.59	55.00	22.19
African American	51.87	55.00	22.76
Posttest truthfulness	51.02	55.00	24.99
Caucasian	55.78	62.00	23.23
African American	47.00	42.50	26.07
Pretest Alcohol	36.73	34.00	28.58
Caucasian	37.26	34.00	28.42
African American	36.28	38.00	29.16
Posttest Alcohol	28.08	25.00	22.49
Caucasian	30.33	26.00	21.98
African American	26.19	25.00	23.08
Pretest Drug	38.29	38.00	27.11
Caucasian	37.26	38.00	25.18
African American	39.16	42.00	29.01
Posttest Drug	28.97	19.00	20.91
Caucasian	23.93	19.00	17.39
African American	33.22	38.00	22.88
Pretest Control	48.69	43.00	29.82
Caucasian	51.89	52.00	31.36
African American	46.00	40.00	28.68
Posttest Control	35.54	27.00	25.09
Caucasian	30.26	17.00	22.82
African American	40.00	43.00	26.38
Pretest Violence	68.59	74.00	23.34
Caucasian	70.26	74.00	23.44
African American	67.19	72.50	23.53
Posttest Violence	55.06	54.00	22.46
Caucasian	53.48	48.00	21.71
African American	54.06	57.00	22.46
Pretest Stress-Coping	52.46	58.00	32.06
Caucasian	51.96	59.50	32.36
African American	52.87	56.00	32.31
Posttest Stress-Coping	44.98	45.00	28.67
Caucasian	43.52	43.00	27.78
African American	46.22	47.00	29.78

NOTE: DVI = Domestic Violence Inventory.

following treatment. Eleven (18.33%) of 60 treatment completers were rearrested compared to a slightly higher rate of 5 rearrests (22.73%) of 22 noncompleters.

Statistical Analyses

Three statistical tests were conducted to assess the extent to which mortality biases may have affected the internal validity of the study, evaluate participants' changes in psychological variables between pretest and posttest, examine differences in change scores by ethnicity, and determine the extent to which rearrests can be predicted from change scores on psychological variables of treatment completion. A Bonferroni's Correction for Inequality was computed to maintain the family-wise alpha level at .05 across the three analyses ($\alpha = .0167$ for each test).

Initially, discriminant function analyses were planned for use in evaluating the research question. However, an initial discriminant function analysis indicated that the assumption of equality of the variance-covariance matrices had been violated with the data. In this situation, logistic regression analyses are an excellent substitution for discriminant function analyses (Wright, 1995).

Two logistic regression analyses were used to analyze pretest DVI scale scores to evaluate mortality effects and to determine the extent to which rearrests within a year can be predicted by the DVI change scores of treatment completers. For the logistic regression analyses, data were examined for the presence of outliers, which can disproportionately influence test results. There were no scores on the scales that were beyond 2.5 standard deviations of the mean.

The third analysis was a repeated measures MANOVA that examined changes in the linear combination of DVI scores for participants between the pretest and posttest and between African American and Caucasian participants (the only two ethnic groups represented by the sample). The assumptions of univariate equality of error variances and multivariate equality of covariance matrices were tested at $\alpha = .01$ through the use of Levene's and Box's M tests, respectively. These analyses were not significant and indicated that the above test assumptions were not violated. The Levene's test yielded significance levels ranging from .084 for the posttest for the Drug scale to a .981 for the Violence posttest scale. The Box's M analysis yielded a score of 104.929 ($F = 1.037, p = .389$).

Evaluation of Mortality Effects

A logistic regression analysis evaluated the extent to which mortality effects (drop-out from treatment) may have affected the results of the study, using the six DVI pretest scale scores to predict treatment completion. The analysis stopped at the third iteration and yielded a -2 log likelihood of 93.592. The omnibus test was not significant ($\chi^2 = 2.403, p = .879$). The six DVI scales did not significantly predict treatment completion, yielding probability levels that ranged from .337 to .835. The classification table provided information about the extent of accuracy in predicting treatment completion. An examination of this table found that all of the treatment completers ($n = 61$) were accurately predicted to have completed treatment (accuracy rate = 100%). However, *all* of the noncompleters were also predicted to complete treatment ($n = 22$), with an accuracy rate of 0%. The combined accuracy rate was 73.5%. This analysis found that treatment completers and noncompleters did not differ significantly in their pretest DVI scale scores, and this finding supported the lack of mortality effects as a likely threat to the internal validity of the study.

Prediction of Rearrest

Change scores on the six DVI scale scores were computed for the participants who completed treatment ($n = 61$), and these scores were used in the second logistic regression to predict rearrest within a period of 1 year. Three treatment completers did not have full sets of pretest and posttest DVI scores and were omitted from this analysis ($n = 58$). The logistic regression terminated at the fourth iteration and had a -2 log likelihood of 51.521. The omnibus test was not significant ($\chi^2 = 1.803, p = .937$). Additionally, none of the six DVI change scores significantly predicted rearrest, with significance levels ranging from .421 to .964. Again, the classification tables yielded the interesting result of predicting at 100% accuracy those who were not rearrested within a period of 1 year ($n = 48$) and 0% accuracy those who *were* rearrested within a year ($n = 10$). These results suggest that one or more factors besides changes on those psychological factors related to domestic violence have more effect than treatment changes in predicting rearrest.

Evaluation of DVI Scale Score Changes Over Time and by Race

The repeated measures MANOVA examined the extent that participants' scale scores changed significantly and in the expected direction between the

pretest and posttest and whether African American and Caucasian participants differed significantly in changes in their scores over time. Tables 2 and 3, respectively, contain descriptive information by race for the pretest and posttest observations and detailed information on the tests of hypotheses. The omnibus test of within-subject change yielded a significant Pillai's Trace score and partial eta squared (η^2) of .397 ($F = 5.695, p < .001$). The partial η^2 indicates that almost 40% of the variance in the linear combination of DVI scale scores is accounted for by time between the pretest to posttest. Additionally, the observed power was very high at .994 for within-subject change.

Table 4 indicates that five of the six DVI scale score changes over time were significant. The Truthfulness scale scores, which measure response validity, did not change significantly from pretest to posttest. The changes in Lethality scale scores were significant ($F = 22.396, p < .001$) and had a Pillai's Trace and η^2 of .282, indicating that about 28% of the variance in change is accounted for by decreases over time on this scale. The changes in Control scale scores were significant and in the hypothesized direction, as well ($F = 10.562, p = .002$). The amount of variance accounted for by these changes was 15.6% ($\eta^2 = .156$). Changes in the Stress-Coping scale were significant ($F = 6.784, p = .012$), with 10.6% of the variance in change being accounted for by this scale. Interestingly, changes in the Alcohol and Drug scales were also significant and in the hypothesized direction, although participants having these problems are referred to other agencies for substance abuse treatment. Changes in the Alcohol scale yielded an F of 8.824 ($p = .004, \eta^2 = .134$), and the Drug scale changes resulted in an F of 7.260 ($p = .009, \eta^2 = .113$). The observed power for all of the univariate examinations was acceptable to very good, with a range from .726 for the Stress-Coping scale to .996 for the Lethality scale.

Changes by race for the omnibus test were not significant, yielding a Pillai's Trace and η^2 of .068 ($F = .628, p = .707$). Additionally, no changes in any of the six DVI scales were significant by race. However, the observed power (of .228) for this analysis was low, and one cannot be confident of this result. It is possible that further research may result in significant main effects by race on changes in the DVI scale scores.

An examination of the within-subject interaction between race and changes over time also was not significant, resulting in a Pillai's Trace and partial η^2 of .130 ($F = 1.299, p = .274$). Similar to the examination of racial differences, the observed power (of .464) for this analysis was low. Only those changes on the Control scale scores approached significance ($F = 3.380, p = .071$), with an observed power of .439. It is possible that further research with increased levels of power would identify a significant

TABLE 3: Logistic Regression Analyses in Evaluating Mortality Effects and Predictions of Rearrest

<i>Variables</i>	<i>Parameter</i>	<i>SE</i>	<i>Wald χ^2</i>	<i>Significance</i>	<i>Odds</i>
Pretest Truthfulness	-.004	.012	.122	.727	.996
Pretest Alcohol	.010	.011	.796	.372	1.010
Pretest Control	.008	.010	.572	.449	1.008
Pretest Drug	-.010	.010	.923	.337	.990
Pretest Lethality	-.003	.014	.043	.835	.997
Pretest Stress-Coping	-.006	.011	.285	.593	.994
Constant	1.475	1.246	1.403	.239	4.372
Change Truthfulness	-.001	.019	.003	.957	.999
Change Alcohol	-.015	.018	.648	.421	.986
Change Control	-.001	.012	.002	.964	.999
Change Drug	-.002	.015	.024	.876	.998
Change Lethality	-.002	.017	.016	.900	.998
Change Stress-Coping	.014	.020	.542	.462	1.014
Constant	-1.562	.456	11.746	.001	.210

TABLE 4: Results of the Repeated Measures Multivariate Analysis of Variance

<i>Results by Levels</i>	<i>F</i>	<i>Partial η^2</i>	<i>Significance</i>	<i>Observed Power</i>
Omnibus tests				
Main effects of race	.628	.068	.707	.228
Within-subjects effects	5.695*	.397	<.001	.994
Interaction of race and within-subjects effects	1.299	.130	.274	.464
<i>F tests of within-subjects change</i>				
Truthfulness	.081	.001	.777	.059
Alcohol	8.824*	.134	.004	.831
Drug	7.260*	.113	.009	.755
Violence	22.396*	.282	<.001	.996
Control	10.562*	.156	.002	.892
Stress	6.784*	.106	.012	.726

* = Significant beyond $\alpha = .0167$.

interaction of race and changes over time for this scale. An examination of the descriptive statistics revealed that Caucasians had relatively higher average scores with a higher standard deviation on Control scores at the pretest and lower average scores and standard deviation on the posttest than African American participants.

DISCUSSION AND APPLICATIONS TO SOCIAL WORK PRACTICE

Question 1: Does the current standardized treatment program for batterers significantly alter psychological variables related to domestic violence, and do changes on these variables predict rearrest at 12 months posttreatment? The findings of this study suggest that court-ordered batterers demonstrate significant changes, in the desired direction, on psychological variables related to domestic violence, as a result of participation in a court-mandated treatment program. However, given the lack of a control group and the low power involved in the analyses, it is unfair to conclude that the positive changes observed on the DVI occurred as a direct result of program participation. Nevertheless, the treatment program appears to be having the desired effect in terms of amending the men's attitudes about women and modifying the behaviors they engage in with their partners. These findings are consistent with those in the literature, which suggest that the majority of batterers who complete treatment are no longer physically abusive toward their partner at the conclusion of the treatment program. However, the findings from this study are even more encouraging because they suggest significant changes across a broader spectrum of violent and coercive behaviors. In essence, if batterers are only modifying their physically abusive behaviors as a result of program completion, it is possible that the lives of their partners and children are not appreciably improved. In brief, if the treatment program only targets "behavioral containment," it is possible that men in treatment reduce the frequency of their episodes of physical aggressiveness but increase their use of emotionally controlling behaviors, becoming in the process more savvy abusers. Fortunately, the results of this study suggest that this is not the case. It appears that for the men in this study, the treatment program successfully targeted more than just physical violence.

Although the findings suggest that the treatment program is having the desired effect on participants, the follow-up data failed to reveal a relationship between changes on these psychological variables and rearrest. There are several possible explanations for this situation. First, and perhaps most important, this study experienced an attrition rate of 27%, which is significantly less than the national average of 40%-60% (Gondolf, 1997). This low attrition rate can be explained by the strong judicial support for the treatment program, which ensures that men failing to either participate actively in or complete treatment eventually will serve some jail time. In brief, when an individual is terminated from the treatment program for nonparticipation or nonattendance, the judge either places the individual in jail or issues a bench warrant for his arrest. This long-standing judicial support of the treatment

program has helped create widespread community awareness of the negative consequences associated with failing to follow through with a judicial referral to the program. Consequently, treatment program participants are aware of the negative consequences associated with noncompliance, and attrition is quite low. In terms of explaining the absence of a significant relationship between change on the psychological variables and rearrest, it is possible that this particular treatment program retains men who would drop out of treatment in another location. Specifically, it is possible that men who, in other locations where judicial support for the treatment program is low, would drop out, continue to attend and participate in the program. In this regard, judicial support for the program may interact with and confound the assessment of rearrest as an indicator of program success.

Another possible explanation for the failure of this study to discover a relationship between significant changes on these psychological variables related to domestic violence and rearrest, is that rearrest is an imperfect measure of domestic violence. As discussed previously, there are many problems related to the accurate measurement of violence in intimate relationships. The most intractable problem is social desirability in that the correct response to questions is obvious (e.g., After completing treatment, do you still slap your partner?) and there is no way to identify those men who may be misrepresenting themselves to avoid a recommendation of continued treatment. Given these measurement limitations, many treatment providers have begun to evaluate rearrest as an indicator of program success, in addition to assessing change on psychological variables. However, there is no way to assess whether men not rearrested are also not being violent in their intimate relationships. Thus, the best alternative currently available is to treat rearrest as an imperfect measure of the recurrence of violent behaviors in intimate relationships and exercise caution in concluding that men not rearrested are also not engaging in violence. On a related note, some treatment providers have begun to interview partners as an indicator of program success. However, given the fact that this study employed a sample of men court mandated into treatment, there was no way to insist on interview compliance after the men completed treatment and were referred back to the courts.

Despite these limitations, the findings suggest that men completing court-mandated treatment for domestic violence offenses are making significant improvements regarding the way they behave in their intimate relationships. However, it is unclear how these within-program changes relate to rearrest. This confusing situation may be related to a lack of connection between significant changes on these psychological variables and long-term behavioral change or the imperfect nature of rearrest as a follow-up tool. Additionally, it is also possible that change in these relationships is incremental and much

like the situation in substance abuse treatment, men experience episodes of relapse on the road to complete recovery. Regardless of the explanation, what remains clear is that measurement problems in the field continue to inhibit meaningful understanding regarding the mechanism of change in BIPs.

Question 2: Does the current standardized treatment program for batterers differentially affect African American and Caucasian batterers? The findings from this study are important because they represent the first empirical research in the field attempting to investigate the differential effect of the standardized treatment program for batterers on outcomes for African American and Caucasian batterers. The findings from this study suggest that the standardized treatment program employed at this agency reaches African American and Caucasian participants equally well. This finding has critically important implications for treatment programs. In brief, although the methodological problems limiting understanding of the efficacy of the standardized cognitive-behavioral treatment approach are well articulated, treatment programs are facing legislative requirements to institutionalize the one size fits all model (Moore et al., 1997). The downside to this lack of connection between research and practice is that intervention programs are being forced to operate within the constraints of the one size fits all model, despite recent research documenting the presence of different batterer subtypes (Hamberger, Lohr, Bonge, & Tolin, 1997; Holtzworth-Munroe, Meehan, Herron, & Stuart, 1999; Holtzworth-Munroe & Stuart, 1994) and without any empirical evidence evaluating the usefulness of this approach for diverse client populations (Bennett & Williams, 2001). Consequently, the findings from this study are important because they provide preliminary empirical evidence for the contention that the standardized cognitive-behavioral treatment program works equally well for African American and Caucasian batterers.

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, despite the Truthfulness scale of the DVI, it is still possible that the men completing the program had a sense of the type of

responses that were expected of them at the conclusion of the treatment program. Consequently, the Truthfulness scale may have only partially corrected for the tendency toward socially desirable responding among this sample.

REFERENCES

- Bennett, L., & Williams, O. (2001). Intervention program for men who batter. In C. Renzetti & J. Edleson (Eds.), *Sourcebook on violence against women* (pp. 261-277). Thousand Oaks, CA: Sage.
- Deavers, L. (1998). *Domestic violence intervention program*. Tuscaloosa, AL: Family Counseling Service.
- Eisikovits, Z., & Edleson, J. (1989). Intervening with men who batter: A critical review of the literature. *Social Service Review, 37*, 385-414.
- Gondolf, E. (1997). Batterer programs: What we know and need to know. *Journal of Interpersonal Violence, 12*, 83-98.
- Hamberger, K., Lohr, J., Bonge, D., & Tolin, D. (1997). An empirical classification of motivations for domestic violence. *Violence Against Women, 3*, 401-423.
- Holtzworth-Munroe, A., Bates, L., Smultzer, N., & Sandin, E. (1997). A brief review of the research on husband violence. *Aggression and Violent Behavior, 2*, 65-99.
- Holtzworth-Munroe, A., Meehan, J., Herron, K., & Stuart, G. (1999). A typology of male batterers: An initial examination. In X. B. Arriaga & S. Oskamp (Eds.), *Violence in intimate relationships* (pp. 45-72). Thousand Oaks, CA: Sage.
- Holtzworth-Munroe, A., & Stuart, G. (1994). Typologies of male batterers: Three subtypes and the differences between them. *Psychological Bulletin, 116*, 476-497.
- 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.
- Moore, K., Greenfield, W., Wilson, M., & Kok, A. (1997). Toward a taxonomy of batterers. *Families in Societies: The Journal of Contemporary Human Services, 78*, 352-359.
- Oliver, W. (2000). The public health and social consequences of Black male violence. *Journal of African-American Men, 5*, 71.
- Plass, P. (1993). African-American family homicide: Patterns in partner, parent, and child victimization, 1985-1987. *Journal of Black Studies, 4*, 515-538.
- Rennison, C., & Welchans, S. (2000). Intimate partner violence. *Bureau of Justice Statistics Report* (NCJ 178247). Washington, DC: U.S. Department of Justice.
- Risk and Needs Assessment, Inc. (1997a). *Domestic Violence Inventory DVI-juvenile and DVI pre-post: DVI, DVI-juvenile and DVI pre-post orientation and training manual*. Phoenix, AZ: Author.
- Risk and Needs Assessment, Inc. (1997b). *Domestic Violence Inventory (DVI): An inventory of scientific findings*. Phoenix, AZ: Author.
- Rosenfeld, B. (1992). Court-ordered treatment of spouse abuse. *Clinical Psychology Review, 12*, 205-226.
- Tolman, R., & Bennett, L. (1990). A review of quantitative research on men who batter. *Journal of Interpersonal Violence, 5*, 87-118.

- Williams, O. (1992). Ethnically sensitive practice to enhance treatment participation of African-American men who batter. *Families in Society: The Journal of Contemporary Human Services*, 73, 588-595.
- Williams, O. (1994). Group work with African-American men who batter: Toward more ethnically sensitive practice. *Journal of Comparative Family Studies*, 25, 91-103.
- Williams, O. (1995). Treatment for African-American men who batter. *CURA Reporter*, 25, 610.
- Williams, O., & Becker, R. (1994). Domestic partner abuse treatment programs and cultural competence: The results of a national survey. *Violence and Victims*, 9, 287-295.
- Wright, R. E. (1995). Logistic regression. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics*. Washington, DC: American Psychological Association.
- Wyatt, G., Axelrod, J., Chin, D., Carmona, J., & Loeb, T. (2000). Examining patterns of vulnerability to domestic violence among African-American women. *Violence Against Women*, 6, 495-515.