

Results From the Brooklyn Domestic Violence Treatment Experiment

By Christopher D. Maxwell, Robert C. Davis, and Bruce G. Taylor

2004
NCJ 199728

Christopher D. Maxwell, Ph.D., is an assistant professor at Michigan State University, School of Criminal Justice; Robert C. Davis, M.A., is a senior research associate at the Vera Institute of Justice; and Bruce G. Taylor, Ph.D., formerly with the U.S. Department of Justice, is a senior research associate at Caliber Associates, Inc.

This study was supported under award 96-IJ-CX-0047 from the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Findings and conclusions of the research reported here are those of the authors and do not reflect the official position or policies of the U.S. Department of Justice, the Vera Institute of Justice, or Michigan State University.

Over the past two decades, society has relied increasingly on criminal justice sanctions to control intimate partner violence. Police departments across the country instituted proarrest policies that were strongly promoted by advocates and were consistent with results from the Minneapolis Domestic Violence Experiment (Buzawa and Buzawa 1996). More recently, prosecutors have also pursued conviction regardless of the victim's desires or willingness to cooperate (Rebovich 1996; Hanna 1996). These policies have led to an expanded pool of batterers that the criminal courts have had to meaningfully sanction. As a result, the courts have become increasingly dependent on batterers' group treatment programs as the sanction of choice (Davis, Smith, and Nichols, 1998).

Group treatment programs for batterers are a popular criminal court sanction for several reasons. First, even in cases involving serious felony assaults, many victims choose to stay with their abusive partners for a variety of personal, family, and economic reasons. These victims are interested in sanctions that offer them long-term safety from violence rather than punishments that jeopardize their partner's ability to earn a living or that may lead to more violence. Second, alternative sanctions commonly used to punish other crimes lack validity in rehabilitating those convicted of intimate partner violence. There is little reason to believe, for example, that fines, community service, or traditional probation will help batterers stop abusing their intimate partners. Therefore, batterer treatment programs theoretically provide both safety for victims who want to stay with their partners and realistic help for batterers who want to address their violent behaviors. However, the limited research conducted during the 1980s that assessed batterer treatment programs could not provide satisfactory answers to those who question whether batterers' programs really provide either victim safety or therapeutic help for the batterers (Davis and Taylor 1999).

Purposes of the Brooklyn Study

Because of the lack of rigorous experimental evaluations of batterer treatment programs, the authors sought and received funding from the National Institute of Justice in 1994 to conduct an experimental evaluation of the Victim Services' Alternatives to Violence (ATV) batterer treatment program in Brooklyn, New York. Until the mid-1990s, only one completed experimental evaluation of a batterers' treatment group was published in the literature. This experiment was conducted by Palmer, Brown, and Barrera in 1992. They randomly assigned batterers to either a 10-session psychoeducational group (combining group discussion with information) or a no-treatment control group. The researchers in that study examined police reports 6 months after treatment and found significantly lower recidivism rates for the treatment group compared with the control group. They also attempted to measure violence through surveys of victims and batterers, but because of low response rates and a small sample size ($n = 59$) the analysis of recidivism based on interview data was not completed.

Beyond the Palmer, Brown, and Barrera (1992) experiment, two other experiments were ongoing while the authors conducted the Brooklyn study. One study by Dunford (2000) randomly assigned Navy personnel who committed domestic violence to one of four groups: group treatment for batterers, couples counseling, a rigorous monitoring program (periodic calls to victims and record checks), or a safety planning program for the victim that was intended to

approximate a no-treatment control group. The second study was implemented by Feder and Forde (2000) in Broward County, Florida. In this study, Feder and Forde had domestic court judges randomly assign male offenders who had been convicted of misdemeanor domestic violence to either a 1-year probation and a Duluth-like batterer treatment group (described below) or a control group that received only 1 year of probation. At this time, neither Dunford nor Feder and Forde have reported comparisons between their groups that have produced a statistically significant difference.

Brooklyn Treatment Group

The authors experimentally evaluated the ATV batterer treatment program, which was based on the Duluth model. The Duluth model mandated 26 weeks of attendance at a weekly 1-hour group meeting. The course was rooted in a feminist perspective that assumes domestic violence is a by-product of conventional male and female sex roles. The curriculum included defining domestic violence, understanding the historical and cultural aspects of domestic abuse, and reviewing criminal/legal issues. Through a combination of instruction and discussion, participants were encouraged to take responsibility for their anger, actions, and reactions. Sessions were conducted in either English or Spanish by two leaders, one male and one female.

At the time the evaluation began accepting clients, the ATV program expanded the number of required treatment hours from 1.5 hours once a week for 12 weeks to 1.5 hours once a week for 26 weeks. The change was made to conform with New York State's guidelines and was in line with national trends. This length, however, was problematic for the Legal Aid attorneys who represented the majority of defendants judged to be indigent in the Brooklyn criminal courts. A meeting with Legal Aid attorneys revealed that their objections to ATV stemmed from the increased time their clients were under court control and the increased session fees their clients had to pay for 26 sessions. To remedy this problem, the authors designed a new, 8-week format with the help of ATV administrators. Clients in this new program would complete 40 hours of group treatment through biweekly 2.5 hour sessions with lower fees per session. This condensed format began after the first 129 participants were already assigned to the 26-week groups. The final 61 ATV clients were assigned to the 8-week group, which allowed the authors to compare length of treatment while holding the dosage (e.g., 40 hours) constant across treatment groups.

With regard to the comparison group, an experimental control group is conceived in several ways. The most common procedure is to compare those assigned to the treatment group (such as batterer treatment) with a group receiving no treatment or intervention (such as probation). This option was not available to the study because probation for those convicted of misdemeanor spouse abuse is rare in New York City. Judges are the ones who mandate batterers to treatment, and completion of the program is normally the only condition in plea arrangements. Therefore, an alternative sanction for the control group was needed that was irrelevant to the offenders' battering problem. The authors believed community service was such a sanction because it required only that offenders work at such tasks as renovating housing units, clearing vacant lots to make way for community gardens, painting senior citizen centers, and cleaning up playgrounds. Criminal justice officials also agreed to use this sanction as an alternative to ATV for men assigned to the control group. Therefore, all participants in the experiment were assigned

either to receive 40 hours of group batterers' treatment or to complete 40 hours of community service.

Design of the Brooklyn Experiment

The study randomly assigned male criminal court defendants who were charged with assaulting their intimate female partners to either 40 hours of batterers' treatment or to 40 hours of community service. The design called for treatment assignments to be made during sentencing, after the judge, the prosecutor, the ATV program, and the defendant all agreed to accept batterers' treatment as a sanction if it was available based on random assignment. This sample framework is somewhat different from other experiments that included all or most batterers sentenced to probation, regardless of a batterer's willingness or unwillingness to enter treatment. These results are not as easy to generalize, therefore, as the results from the other three experiments on batterer treatment programs. Because everyone in the sample agreed to treatment, however, the current study presumably included only batterers who were motivated to stop battering. This is a key point, because it has been argued that treatment cannot work for individuals who are in treatment against their will.

To measure recidivism, data from several sources were collected to develop multiple indicators of new violence by the batterer against the victim. These included arrest reports, crime incident complaints (which may or may not result in an arrest), victim reports of violence by the batterer, and batterer reports of assaulting the victim. Batterer and victim reports were gathered from interviews that occurred at the time of sentencing, 6 months after sentencing, and 12 months after sentencing. To assess the frequency and severity of violence through the interviews, the study employed Harrell's (1991) revision of the Conflict Tactics Scale. Harrell's scale measures the frequency of 11 violent acts: (1) Forced you to have sex; (2) choked or strangled you; (3) threatened to kill you; (4) beat you up; (5) threatened you with a knife, gun, or other weapon; (6) used a knife, gun, or other weapon against you; (7) threw something at you; (8) pushed, grabbed, or shoved you; (9) slapped or spanked you with an open hand; (10) kicked, bit, or hit you with a fist; and (11) hit or tried to hit you with something.

In the outcome models, the authors examined the combined frequency of all 11 violent acts reported by the victim at the 6- and 12-month intervals. The reference period for the scale was the previous 2 months because it was believed that treatment would take some time to have an effect. Asking victims to report at the 6-month interval about the entire period could include reports of violent incidents committed shortly after batterers were assigned to treatment. The 2-month reference period would increase the likelihood that reported incidents occurred after the batterers completed approximately 40 hours of treatment. Unfortunately, as in other studies, only the females identified as the victim in the triggering court case were interviewed and not new female intimate partners.

In addition to interviews, the authors also gathered data from the computerized records of the New York City Criminal Justice Agency and the New York City Police Department for incidents that occurred during the study period or arrest reports filed against the batterer during the study period. When new incidents were found, the arrest date and charge were recorded. In addition,

the district attorney's computer database was searched using the docket number to determine whether the victim in the new incident was also the victim in the original incident. To reduce measurement error and double counting, the arrest reports and the police complaint data were merged into one measure that captured the number of documented criminal justice incidents (e.g., arrests or crime complaints) involving both the defendant and victim after treatment was assigned.

Findings

The study sample contained 376 adult males, which was approximately 3.4 percent of the estimated 11,000 domestic violence defendants adjudicated in the Brooklyn Criminal Courts within the study's intake period (February 19, 1995 to March 1, 1996). Nearly two-thirds (64 percent) of defendants in the sample were charged with third-degree assault (a Class A misdemeanor), while the remaining third were charged with felonious assault (19 percent), violating restraining orders, menacing, harassment, or other charges. Conditional discharge was the most common disposition for individuals in the sample (76 percent), followed by cases adjourned in contemplation of dismissal (22 percent) and probation (2 percent). A typical subject in this sample was a 30-year-old African-American male (40 percent) who had no prior criminal history (63 percent), no education beyond a high school diploma (7 percent), some but not consistent employment (63 percent), and a personal income of about \$16,000 per year, who was married to (41 percent) or living with his victim (20 percent) at the time of arrest.

Table 1 sets forth the results from both the 6- and 12-month victim interviews and the 12-month followup using police records. The primary outcome measure from the two victim interviews was the frequency of reported aggressive incidents by the suspect against the victim. The study's analysis finds that at both 6 and 12 months, the average frequency of incidents reported by victims in both the 8- and 26-week treatment groups was substantially reduced compared with victims in the control group. At the 6-month interviews, the frequency of victim-reported incidents dropped by 33 percent for the 8-week group and 36 percent for the 26-week group. At 12 months, the reductions were 46 percent for the 8-week group and 21 percent for the 26-week group.

Table 1. Percent Reductions in the Rate of Incidents, by Assigned Treatment Group

	Victim Interviews (Frequency)		Police Reports and Arrest	
	6 months	12 months	12-month frequency	Time to first incident
Control group compared with 8-week group	67.4%	54.4%	20.9%	14.2%
26-week group	64.0%	79.3%	43.4% *	52.2% *

Note: Estimated differences in the rates of incidents are based on multivariate models that also control for the batterer's age, marital status, ethnicity, employment, and criminal history record.

* $p < .01$.

Generalizing these results beyond this sample, the authors stress, requires caution for two reasons. First, none of the sizable reductions in aggression reached the generally acceptable levels of statistical significance ($p < .05$). In large part, this was due to the second reason: not all victims were interviewed at either the 6- or 12-month followup periods. The completion rate for the victim surveys was 48 percent for the first followup interview and 50 percent for the second followup interview. Overall, the authors were unable to contact 131 victims (35 percent of the total sample of 376 victims) during the entire followup period. In many cases, they found out from other sources that the victims had moved. Fortunately, the interview completion rates were not significantly different by assigned treatment groups at either the 6- or 12-month followups. Furthermore, no significant differences in interview completion rates were seen across several demographic measures, except for the victim's ethnicity. The authors had better success interviewing Hispanic victims than African-American victims for the 6-month interviews, but they had a higher completion rate for African-American victims than victims from one of the "other" racial groups (mostly whites and Asians) at both the 6- and 12-month intervals. It is not clear how these differences may have affected the results.

With regard to the outcomes based on official police reports, both the 8- and 26-week groups had substantially lower rates of failure than the control group. Over a 12-month period, the 8-week group had 20 percent fewer total incidents than the control group and the 26-week group had 43 percent fewer total incidents. Similar results were also found by examining the time to the first new incident recorded by the police. Compared with the control group, those in the 8-week group were 14 percent less likely to have a new incident any day after treatment was assigned, and the 26-week group was 52 percent less likely. Again, the authors urge caution in generalizing these positive results because only those reductions reported for the 26-week group were statistically different from the control group.

Discussion

The findings from the experimental evaluations of the ATV batterers' treatment program provide useful information and hypotheses for future researchers and practitioners. First, regardless of the source of outcome data, the authors found consistent reductions in the rate of violence by the batterers against their victim who were assigned to the ATV treatment program. Second, in seven of the eight comparisons, the largest reductions were found among batterers who attended the 26-week treatment program. The authors are guarded, however, about claiming unequivocally that treatment worked better than community service at reducing violence or that longer treatment is better than shorter treatment. Only two of the eight comparisons reached statistical significance, and the authors were only able to interview about 65 percent of the victims after treatment. Furthermore, they are not sure whether longer treatment necessarily led to greater reductions in violence or whether violence was reduced only because longer treatment provided greater supervision of the clients in the community. More research is necessary to replicate these positive findings and to explore whether treatment or supervision was the mechanism behind the apparent additional positive effect from the longer treatment.

Implications for Researchers

Future research projects should consider several important lessons from this study. First, as recognized by Fagan (1996) and others, randomized experiments should be the design of choice when asking questions about which alternative batterer sanction is more effective at increasing victim safety. Random assignment of offenders when applied by the judiciary is difficult to implement; however, the study's research as well as Feder and Forde's (2000) shows that an experimental design is still a realistic choice in jurisdictions where treatment is not yet mandated by legislation. Unfortunately, the opportunity to conduct further experimentation is becoming less available as more government organizations institute standards and mandates that reduce treatment options (see Minnesota Center Against Violence and Abuse 2001). An example of this policy change is the Michigan Governor's Task Force on Batterer Interventions Standards' 1998 recommendation that batterers attend 52 or more treatment sessions, while also acknowledging that "research does not necessarily point to a particular length."

Second, the research community should work toward measures and followup intervals that are standardized so that data and results can be compared across studies. For instance, researchers ought to include both victim and batterer interviews and collect documentation to measure complaints made to the courts by the victim as well as police incident and arrest reports involving the batterer, as was done in the National Institute of Justice's Spouse Abuse Replication Project (SARP) studies. In addition, batterers ought to be tracked for at least 18 months, but preferably 2 years, with interview intervals no longer than 6 months. The short-term measures are needed to assess immediate treatment effects and the longer-term followups are needed to determine whether treatment leads to permanent change or is transitory. The use of both short-term and long-term measures is especially important in light of some results from SARP that showed that arrest may have large positive effects early, but these effects are minimal 1 year after arrest (Parmley and Maxwell 2000).

Implications for Practitioners

Practitioners can also ask questions about the effectiveness of their local programs and not just assume that something they may have called treatment is helpful. They should ask the research community for explicit evidence about the extent to which research findings like those reported in this and other papers in this series can be generalized to their locality. When asking for evidence, practitioners should also pay particular attention to the nature of the sample of batterers who participated in a research study. Are the batterers under court supervision? Do they have extensive prior criminal histories? Do defendants have a chance to volunteer for treatment or are they sent to treatment regardless of their willingness to participate? Another important issue is the criminal justice context in which treatment studies are set. Unfortunately, because of the small number of studies little is known about how treatment program effectiveness varies with local court practices, linkages between agencies, sanctions for noncompliance, and so forth.

There are parallels between batterer treatment literature today and the literature on the rehabilitation of criminal offenders 20 years ago. In each literature, the problem is not too few studies, but a paucity of sophisticated research. Recommendations made years ago by the National Academy of Sciences (Martin, Sechrest, and Redner, 1981) for agreement on outcome measures and randomized experiments in rehabilitation are just as relevant today for batterer treatment programs. The evolution in sophistication of batterer treatment studies is encouraging. Using randomized experiments and other designs that have a high degree of internal validity, researchers should soon be able to estimate the extent to which batterer treatment reduces aggression and violence and to specify which program models are relatively more effective.

References

- Buzawa, E.S., and Buzawa, C.G. (1996). *Domestic Violence: The Criminal Justice Response*, 2d ed. Newbury Park, CA: Sage Publications.
- Davis, R.C., Smith, B.E., and Nickles, L.B. (1998). "The Deterrent Effect of Prosecuting Domestic Violence Misdemeanors." *Crime and Delinquency*, 3(44), July: 434–442.
- Davis, R.C., and Taylor, B.G. (1999). "Does Batterer Treatment Reduce Violence? A Synthesis of the Literature." *Women and Criminal Justice*, 10(2): 61–91.
- Dunford, F.W. (2000). "San Diego Navy Experiment: An Assessment of Interventions for Men Who Assault Their Wives." *Journal of Consulting and Clinical Psychology*, 68(3), June: 468–476.
- Fagan, J.A. (1996). *The Criminalization of Domestic Violence: Promises and Limits*. Presentation at the Conference on Criminal Justice Research and Evaluation. Washington, DC: U.S. Department of Justice, National Institute of Justice, NCJ 157641.
- Feder, L., and Forde, D.R. (2000). Test of the Efficacy of Court-Mandated Counseling for Domestic Violence Offenders: The Broward Experiment. Final report submitted to the U.S. Department of Justice, National Institute of Justice, NCJ 184752.
- Hanna, C. (1996). "No Right to Choose: Mandated Victim Participation in Domestic Violence Prosecutions." *Harvard Law Review*, 109(8), 1849–1910.
- Harrell, A. (1991). *Evaluation of Court-Ordered Treatment of Domestic Violence Offenders*. Supported by the State Justice Institute (#90-121-E-089), under subcontract from the Institute for Social Analysis. Washington, DC: The Urban Institute.
- Martin, S.E., Sechrest, L.B., and Redner, R., eds. (1981). *New Directions in the Rehabilitation of Criminal Offenders*. Panel on Research on Rehabilitative Techniques, 498. Washington, DC: National Academy Press.

Michigan Governor's Task Force on Batterer Intervention Standards. (1998). "Batterer Intervention Standards for the State of Michigan." (<http://www.biscmi.org/>)

Minnesota Center Against Violence and Abuse. (2001, 9 September). "A Review of Standards for Batterer Intervention Programs." In Violence Against Women Online Resources. (<http://www.vaw.umn.edu/Vawnet/standard.htm> Minneapolis, MN.) Accessed 09/10/02.

Palmer, S., R. Brown, and M. Barrera. (1992). "Group Treatment Program for Abusive Husbands: Long-Term Evaluation." *American Orthopsychiatric Association, Inc.*, 62(2), 276–283.

Parmley, A.M., and C. Maxwell. (2000). "Use of Multiple Outcome Measures to Assess the Impact of Arrest on Violence Against Female Intimate Partners: The Milwaukee Domestic Violence Experiment Revisited." Paper presented at the National Institute of Justice Research Presentation Brown Bag Series, Washington, DC.

Rebovich, D.J. (1996). "Prosecution Response to Domestic Violence: Results of a Survey of Large Jurisdictions." In E.S. Buzawa and C.G. Buzawa, eds., *Do Arrest and Restraining Orders Work?* Thousand Oaks, CA: Sage Publications: 176–191.