



What is the Evidence for the Effectiveness of Perpetrator Programmes? *

*Lesley Laing
Clearinghouse Director*

Introduction

Developing a response to men who use violence and abuse in their intimate relationships is one of the most controversial issues facing the field. Debate tends to centre most strongly around the issue of treatment/education programmes for perpetrators, often with people adopting polarised positions either “for” or “against” the development of such programmes. This paper will review the evidence for the effectiveness of programmes for perpetrators of domestic violence drawing on recent international research.

Because of the controversy surrounding responses to perpetrators of domestic violence, the question of the effectiveness of perpetrator programmes has been extensively canvassed. Yet it is clear that there is no easy answer to the question as to whether or not perpetrator programmes “work” and that teasing out the complexities which underlie this apparently simple question is a challenging endeavour.

The earliest efforts to evaluate outcomes of perpetrator programmes were plagued by an array of methodological problems, including, for example:

- Lack of agreement about what constitutes “success” or “effectiveness” (Muller 1997). Is it, for example, a total cessation of all forms of abusive, coercive

* Copyright © Australian Domestic and Family Violence Clearinghouse 2003

and controlling behaviour; cessation of physical abuse; or a reduction in the amount of physical abuse? (Edleson 1995; Gondolf 1997a)

- Small sample sizes. This is common in the Australian research (e.g. urbis keys young 2001a).
- Disagreement over what constitutes an adequate follow-up period.
- Reliance on self-reports of change by the men or rearrest records as outcome measures, both of which under-count re-offending (Palmer, Brown, & Barrera 1992).
- Low response rates in follow-up because of high rates of programme drop out and difficulty in tracing participants over extended follow-up periods.
- Difficulty of involving partners in follow-up, for example because of the risks of infringing on their privacy, jeopardising their safety, or other reasons.
- Inclusion of only initial partners, leaving unanswered the question of whether men simply transfer their abusive behaviour to new partners (Gondolf 2002).
- No control or comparison group used so that any changes identified cannot necessarily be attributed to the programme.
- Evaluations conducted by staff with a bias towards demonstrating “success” (Gondolf 2002).
- Problems in operationalising outcomes for women which aim to increase their safety and well being (Austin & Dankwort 1999).

As the field has begun to address some of these issues, other areas of debate have emerged, such as the debate about which type of research design can provide the best evidence for evaluating perpetrator programme outcomes. There are some who contend, for example, that nothing less than evidence of effectiveness using an experimental research design is required to validate their operation. Experimental designs randomly assign participants to two groups – a control group, which does not receive the intervention being tested, and a treatment group which does receive it. The results for the two groups are then compared, to establish if the intervention has had an impact.

Another type of research design, quasi-experimental, uses a comparison group, but without random assignment. In perpetrator programme evaluations, for example, the comparison group may comprise programme “dropouts”, rather than men randomly assigned to a “no intervention” group. The obvious problem here is that the comparison groups may be different - perhaps the men who drop out are more dangerous and disturbed, hence any differences between the groups are due to the characteristics of the two groups of participants rather than to the impact of the programme. Statistical techniques are employed to attempt to control for such differences between the groups.

While some argue that it is unethical to legally mandate treatment that has not been proved effective using an experimental design, others argue that in a complex area such as domestic violence, where programmes for perpetrators are but one component of a coordinated system of intervention, experiments are “artificial” and “naïve” about programme context (Gondolf 2001).

Evaluations which use an experimental research design

There have now been 4 studies of perpetrator programme outcomes using an experimental design.

Palmer, Brown and Barrera (1992) compared 59 court ordered perpetrators who participated under probation in a 10 week psycho-educational, client-centred group programme with a control group assigned randomly to probation only, twelve months after programme completion. Although the research design aimed to collect outcome data from the men, their partners and from police reports, the low response rates, particularly for the women, led to reliance on police data. Recidivism, measured by police records, was significantly higher for the control group than for the programme group. However, this study suffered from small numbers and reliance on police data.

Using random group assignment, **Dunford** (2000) compared four, 12 month interventions with naval base personnel: a men’s cognitive behavioural group; a couples’ counselling group using a cognitive behavioural approach; rigorous monitoring of the men; and no intervention with the men (although safety planning

was implemented with the women partners). This study found no difference in outcomes between the treatment conditions during a year-long follow-up on four outcome measures (self report by the men and their partners; the modified Conflict Tactics Scale; police and court reports; and date of the first repeat case of spouse abuse.) Dunford (2000, p. 475) concluded that: 'The interventions of the cognitive-behavioral model failed to produce meaningful changes in the behaviour they were designed to impact.'

In many ways, the men in this study were not typical of other men referred to perpetrator programmes: all were married, were much younger, and had shorter abuse histories and less drug and alcohol problems than non-military perpetrators. Dunford argues that the failure to demonstrate the effectiveness of cognitive behavioural group intervention in this "optimal" setting sounds a note of caution for programmes run under more typical community settings.

Others have interpreted these results differently. Bennett and Williams (2001) note that the overall recidivism in the study was low (30 per cent on the women's report and 4 per cent by arrest) and attributes this to the fact that the men, by virtue of their being in the navy, were subject to more extensive supervision and decisive sanctions (loss of job and housing) than are men in the general community. Hence they argue that this study:

...while questionable as an indicator of batterer program effectiveness, is nevertheless useful as an indicator of coordinated community intervention...If communities take a proactive response to domestic violence, including assertive probation work, sanctions for non-compliance, victim safety monitoring, and batterer intervention programs, they will probably reduce the incidence of repeat violence. (Bennett & Williams 2001, p. 4)

In **New York**, 376 men were randomly assigned to a six-month perpetrator programme (treatment group) or to 6 months community service (Taylor et al. 2001). During the study, some men in the treatment group were assigned to an alternative treatment programme of two month's duration, with two sessions a week rather than one. The control group comprised the men assigned to community service. Outcomes were measured by interviewing the women partners of the men at

sentencing, and again six and 12 months later and by police record search 12 months after sentencing. Completion rates for victim surveys, however, were low: 51 per cent at intake, 48 percent for the second, and 50 per cent for the third interview. Thirty five per cent of victims were never able to be contacted at any time during the follow-up period.

The evaluation found that the men in the longer (six-month programme) were significantly less likely to be arrested for domestic violence than the men in the control (community service) group and than men in the shorter treatment group. This treatment effect did not diminish significantly between the six and 12-month follow-up periods. However, when the researchers looked at the *victim reports of new incidents*, there were no statistically significant differences between the groups. So this is a mixed but somewhat positive result, although criminal justice recidivism data tends to report lower rates of re-offending than do victims.

This study also encountered difficulties in implementing random assignment: judges over-rode the random assignment to community services in 14 per cent of cases, and the shorter programme was introduced to appease defence lawyers. It has also been critiqued because the sampling frame included only cases in which the prosecutor, defendant and judge agreed that treatment was appropriate. This recruitment process probably filtered out men who were unmotivated to attend a programme.

In **Broward County, Florida**, all men (404) convicted of misdemeanour domestic violence in two specialist domestic violence courts over a five month period were randomly assigned to either one year's probation with 26 weeks' court mandated perpetrator counselling, or to one year's probation, with no perpetrator programme (Feder & Dugan 2002). Data was collected from batterers, victims and criminal justice records. The men and their initial partners were interviewed at adjudication, six months later, and the women 12 months after adjudication. Tools measuring attitudes towards women and wife abuse, the criminalisation of domestic violence, responsibility for domestic violence, the likelihood of hitting/being hit again and the revised Conflict Tactics Scale, were administered to the men and women. Probation and arrest records were also checked at 12 months post adjudication.

No evidence was found that the men in the programme did any better either behaviourally, or in attitudinal change, than the men who only received probation.

The results show that, in this county, there were no clear and demonstrable positive effects of this court-mandated SAAP [spouse abuse abatement] program on the offenders' attitudes, beliefs, and behaviours. Analyses failed to uncover differences between control and experimental subjects in their likelihood of reoffending and being arrested during the follow-up period. (Feder & Dugan 2002, p. 371)

The researchers encountered considerable opposition in implementing this study, one consequence of which was limited access to victims. However, the researchers argue that the low victim response rate was compensated for by the use of multiple sources of data. They conclude:

...an unquestioning acceptance of domestic violence batterers' intervention needs to be challenged. (Feder & Dugan 2002, p. 372)

Summary – experimental studies

It was hoped that outcome studies using an experimental design would unequivocally answer the question as to whether or not perpetrator programmes are effective. However, reviews of the results of these experimental studies have led to varying conclusions by different researchers. Feder and Dugan (2002) and Dunford (2001) see their studies as raising serious questions about the effectiveness of perpetrator intervention programmes. On the other hand, Bennett and Williams (2001a) conclude that the evidence of programme effectiveness from the experimental studies to date is inconclusive. Further, they argue that it is both difficult and undesirable to try to distinguish the effects of perpetrator programmes from the impact of the co-ordinated responses within which they are located. Gondolf makes a similar point in his review of the experimental studies:

The evaluations together suggest that the effectiveness of batterer programs alone is not readily apparent or rather weak and that claims of overwhelming success should be regarded with suspicion. They also indirectly imply that more attention needs to be given to program context. The encompassing

intervention system of arrest, court action, victim services, and probation monitoring may substantially affect program success. (Gondolf 2001, p. 87)

Recent Quasi-Experimental Evaluation Studies

Russell and Rebecca Dobash and colleagues compared the effects of two Scottish court mandated men's programmes with alternative, traditional forms of criminal justice dispositions such as fines, probation or imprisonment (Dobash et al. 2000).

The study compared two naturally occurring groups: the "Men's Program Group" comprising men who were sentenced to and who completed one of two abuser groups as a condition of their probation; and the "Other Criminal Justice Group" (Other CJ) comprising men sentenced by the courts for a domestic violence offence, who received some other form of sanction (Dobash et al. 2000, p. 72). Data was gathered via an initial in-depth interview with the men and women and by postal questionnaire at two follow-up times, three and twelve months after initial contact.

The study found that very few men in either the programme group or the comparison group were charged with further violence towards their partners, indicating no difference on the outcome measure of criminal recidivism. In contrast to the recidivism data, the women's reports revealed much higher rates of re-abuse for both groups. However, the programme group were more successful at reducing their violent and controlling behaviour.

- At the twelve-months follow-up, 75 per cent of the comparison group had re-assaulted, compared to 33 per cent of the 'programme' group;
- Further, at the 12-month point, 37 per cent of the women partners of men in the comparison group reported frequent violence compared with seven per cent of women partners of men in the programme group (Dobash & Dobash 1997);
- Women living with men in the programme group also reported significant reductions in controlling behaviours;
- With respect to findings about quality of life, women partners of men in the programme group were: '...much more likely than women in the Other CJ

group to say they were happy, more relaxed and less frightened than before the intervention...An overwhelming majority of women in the Programme group also indicated that their partner was less likely to try to restrict their lives and more likely to take responsibility for their violence.’ (Dobash & Dobash 1997, p. 251)

An important feature of this study was its use of the women’s reports as its main indicator of outcome and the development of new tools which attempted to more adequately capture all aspects of abuse – the violence, its impact, and the core dynamic of coercive control. Dobash et al. (1999) assert that their results point to the need for mandated treatment, since the programmes studied suffered little attrition, a common problem for many programmes. However, numbers in the study were quite small.

In the USA, a large multi-site evaluation using a quasi-experimental design was funded by the National Centre for Injury Prevention and Control, Centres for Disease Control and Prevention (Gondolf 2002). In designing the study, the research team aimed to address many of the methodological problems identified in earlier outcome studies and to be sensitive to the concerns of those working with perpetrators and with abused women. Key to the approach adopted was the recognition that perpetrator programmes do not operate in isolation, but rather, are embedded in a broader context:

...batterer intervention programs are part of a broader intervention system. They depend on — or at least are related to — arrest practices, court procedures, probation supervision, battered-women’s services, and other community services. (Gondolf 2002, p. 2)

Hence, the evaluation is of the “batterer intervention system” (Gondolf 2002, p. 2), rather than of batterer “programmes”.

The Gondolf research design comprised several innovations, compared to previous studies. These included:

- Use of “naturalistic comparison”. Four programmes were examined as conducted within their communities. Site visits by members of the research team explored the programme approach and contextual changes over time.
- A multi-site study. Research participants were recruited from four sites in different geographical areas and with different client racial and ethnic mixes – Pittsburgh, Houston, Dallas and Denver.
- Sites were selected with different formats and lengths. They ranged from a three-month, pre-trial programme to a nine month, post-adjudication programme with additional specialist components such as personality assessment at intake, alcohol treatment and a women’s services co-ordinator. However, the four programmes had a common structure and format, meeting criteria for being “well established”: compliance with the relevant state standards, collaboration with the battered women’s services in their communities, use of a cognitive-behavioural approach; and being operational for five years or more with at least 40-50 referrals per month.
- Uniform intake procedures were implemented at each site. Each site used the same assessment tools and compiled similar records about issues such as attendance and dropouts, contributing to the large sample in the study.
- Longitudinal follow-up. Phone interviews were conducted every three months with the men and their initial partners, initially for 15 months from programme intake, with a subsequent extension to four years.
- New partners were included in the follow-up where they were identified.
- Both qualitative and quantitative data were collected.
- The main measure of reassault was the women’s reports. These reports were further supplemented by arrest records, the men’s reports and medical records of a sub group of the women.
- Intervening variables were assessed, including not living together, legal action, counselling, drug and alcohol treatment, economic assistance, informal help.

- Large sample size (840) compared to many previous studies. Of this sample, the majority were court mandated (82 per cent), rather than voluntary (Gondolf 2002, pp. 65-68).

In this study, the main outcome measure selected was reassault. The *cumulative* reassault rates were calculated for all court mandated men, whether they completed the programme or dropped out. Based on the women's reports,

- 2 per cent reassaulted during the first, 15-month period of the follow-up.
- This increased to 37 per cent at the 30-months follow-up, and to 42 per cent by the 48-month follow-up.

When adjusted for under-reporting by the women, using men's reports and arrest data, the percentages who reassaulted at the three points in the follow-up were higher: 40 per cent at 15 months; 45 per cent at 30 months; and 48 per cent at 48 months (Gondolf 2002).

From the cumulative reassault rates, it can be seen that almost half the sample committed at least one reassault. However, a different slant on the outcomes can be seen when the *trends in reassault over time* are identified. This different view is possible because of the longitudinal research design:

Overall, the cumulative reassault rate showed that a substantial portion of the men reassaulted, but the trend of the reassault revealed a de-escalation over time and a sustained cessation of violence for the vast majority of men. (Gondolf 2002, p. 113)

The results of this study suggest that the majority of men in the programmes 'eventually do stop their violence, apparently for long periods of time.' (Gondolf 2002, p. 123)

- At 30 months, 'more than 80% of the men had been violence-free for at least a full year' (p. 122)
- At 48 months, 'nearly 90% of the batterers had reportedly not reassaulted a partner in the previous year; and three quarters had not been assaultive for the previous 2.5 years' (Gondolf 2002, p. 122).

The different findings from looking at cumulative rates of reassault and trends in reassault over time initially appear contradictory. They suggest that a high proportion of men commit at least one reassault, but that much of this re-offending occurs early in their involvement with the programme. This is understandable given the serious nature of the problem which perpetrator programmes are addressing, at a point at which the men have had only minimal exposure to the programme.

The finding that the risk of reassault is highest when men are first in the programme, has implications for the intensity of the programme offered, the intensity of the legal supervision of the men, and the level of support offered to women in the early stages of perpetrator programmes. Gondolf suggests, for example, that men might be required to attend three or four times weekly for the first month or so.

The study identified a group of men who reassaulted repeatedly and at dangerous levels throughout the entire follow-up period. The research team devoted considerable effort to establishing whether this dangerous group of men could be identified, by looking at the utility of commonly used risk assessment tools, personality profiles, and batterer typologies. However, the predicative power of these tools was found to be weak.

The repeat reassaulters—the most dangerous men in our evaluation—were not as readily distinguishable as we had thought and hoped. Remarkably, the majority appeared to be acceptable candidates for conventional batterer counseling. Most did not appear pathological, and only a small portion appeared to be truly psychopathic. The mode of violence among these men was also not substantially different from those of other men. (Gondolf 2002, p. 191)

What emerges about this group of men is that they were able to continue to reassault and inflict serious harm because of a *system's failure*: they reassaulted and faced no consequences. Based on these findings, Gondolf suggests that, in contrast to current practice, where risk assessment is undertaken at the beginning of intervention, risk assessment needs to continue throughout the intervention period.

A commonly voiced concern about perpetrator programmes is that the men may simply substitute and increase other, non-physical forms of abuse to maintain their

coercive control over their partner (Gondolf 1997a). In this study, the women were asked about these other forms of abuse. Based on the women's reports, all forms of non-physical abuse reduced over time, along with reassault. However, they remained at relatively high levels, and Gondolf recommends that programmes develop better methods of addressing these other forms of abuse.

The research also aimed to address a common criticism of perpetrator programme evaluations: that they pay insufficient attention to the subjective experience of the victims. Is a programme successful, for example, if the man's violent and abusive behaviour decreases or stops, yet the woman continues to live in fear? This issue was addressed by exploring the women's subjective view of their own well being using a "quality of life" inventory. This tool was developed through consultation with women's advocates and women who had experienced domestic violence (Gondolf 1997b).

The research found that: 'The majority of women in our evaluation...indicated that they were "better off" or "felt safer" and their experience corresponds with the decreases in reassault and abuse.' (Gondolf 2002, p. 127) At the 48-month follow-up, 85 per cent of initial partners said that they felt "very safe", and 84 per cent indicated that it was "very unlikely" that their partners would hit them.

Gondolf also notes that, corresponding with the group of men who assaulted at frequent and harmful levels, there was a group of women who were "worse off" (12 per cent at 15 months and 6 per cent at 30 months) following their partner's participation in a perpetrator programme:

These women told the kind of horror stories that have given batterer programs a bad name. Their experiences are the ones that raise concerns about batterer programs among some battered-women's advocates and confront batterer program staff with the limitations of their efforts. (Gondolf 2002, p. 128)

It is frequently assumed that men who attend programmes voluntarily will be more motivated to change. However, in this study, the voluntary participants (approximately 18 per cent of the sample) were almost 'twice as likely to drop out as the court-referred men (61% vs. 33%), and they reassaulted their partners at a

significantly higher rate at the 15 month follow-up (44% vs. 29%).' (Gondolf 2002, p. 119).

One interesting and unexpected finding of this study was that the outcomes were relatively equivalent across the four programme sites included in the study, despite the fact that the programmes varied in length (from three to nine months) and that the longest programme offered an additional array of specialist service components. One possible explanation is that this finding could be due to the quality of programme implementation at the different sites. However, Gondolf (1999b) suggests that the equivalent success of the shortest programme (Pittsburgh) may be due to characteristics of the broader system in which it is located. This is a three-month, pre-trial programme, with swift court involvement and regular court review: 'In sum, the men were held more quickly and decisively accountable for their behavior.' (Gondolf 2002, p.203)

In summing up the findings of this comprehensive research project, Gondolf (2002, p. 199) concludes that the major implication of the study is that: *The system matters*.

Conclusion

Recent international studies address many of the failings of an earlier generation of research, although many challenges lie ahead, such as:

- Better measuring outcomes for the women partners of men who attend programmes. Beyond checking the accuracy of the men's self-reports of change, we need to know in what ways women may benefit or be adversely affected by their partner's participation in a programme.
- Developing programmes which meet the needs of Indigenous communities and which reflect the cultural diversity of the community.

References

Austin, J. B., & Dankwort, J. 1999, 'The impact of a batterers' program on battered women, *Violence Against Women*, vol. 5, no. 1, pp. 25-42.

- Bennett, L., & Williams, O. 2001, *Controversies and Recent Studies of Batterer Intervention Program Effectiveness*. VAWnet Applied Research Forum, Available: http://www.vawnet.org/VNL.2/Resources/Research/AR_bip.pdf
- Bennett, L. & Williams, O. 2001a, *In Brief: Controversies and Recent Studies of Batterer Intervention Program Effectiveness*. VAWnet Applied Research Forum. Available: www.vawnet.org/VNL.2/Resources/Research/AR_bip.pdf
- Dobash, R. E., Dobash, R., P., Cavanagh, K. & Lewis, R. (eds.) 2000, *Changing Violent Men*, Sage Publications, Thousand Oaks.
- Dobash, R. E., & Dobash, R. P. 1997, 'Men's violence and programs focused on change', *Current Issues in Criminal Justice*, vol. 8, no. 3, pp. 243-262.
- Dobash, R. P., Dobash, R. E., Cavanagh, J., & Lewis, R. 1999, 'A research evaluation of British programmes for violent men', *Journal of Social Policy*, vol. 28, no. 2, pp. 205-233.
- Dunford, F. W. 2000, 'The San Diego Navy experiment: an assessment of interventions for men who assault their wives', *Journal of Consulting & Clinical Psychology*, vol. 68, no. 3, pp. 468-476.
- Edleson, J. L. 1995, *Do batterers' programs work?* Available: www.mincava.umn.edu/papers/battrx.htm [2001, 3 May].
- Feder, L., & Dugan, L. 2002, 'A test of the efficacy of court-mandated counseling for domestic violence offenders: the Broward experiment', *Justice Quarterly*, vol. 19, no. 2, pp. 343-375.
- Gondolf, E. W. 1997a. 'Batterer programs - what we know and need to know', *Journal of Interpersonal Violence*, vol. 12, no. 1, pp. 83-98.
- Gondolf, E. W. 1997b, 'Patterns of reassault in batterer programs', *Violence & Victims*, vol. 12, no. 4, pp. 373-387.
- Gondolf, E. W. 2001, 'Limitations of experimental evaluation of batterer programs', *Trauma Violence & Abuse*, vol. 2, no. 1, pp. 79-88.

- Gondolf, E. W. 2002, *Batterer Intervention Systems: Issues, Outcomes and Recommendations*, Sage Publications, Thousand Oaks:
- Muller, D. 1997, 'Aspects involved with evaluating family violence programs', *Domestic Violence and Incest Resource Centre Newsletter*, no. 1, pp. 11-17.
- Palmer, S. E., Brown, R. A. & Barrera, M. E. 1992, 'Group treatment for abusive husbands: long-term evaluation', *American Journal Of Orthopsychiatry*, vol. 62, no. 2, pp. 276-283.
- Taylor, B. G., Davis, R. C. & Maxwell, C. D. 2001, 'The effects of a group batterer treatment program: a randomised experiment in Brooklyn', *Justice Quarterly*, vol. 18, no. 1, pp. 171-201.
- urbis keys young 2001, *Evaluation of the Learning To Relate Without Violence And Abuse Program: Final Report*, Department of Justice and Community Safety, Canberra.