

To BIP, or not to BIP?

Paper prepared at the request of Hon. Joyce A. Wheeler

By Robert Moyer, Ph.D.ⁱ

Presented to York/Springvale (ME) Domestic Violence Case Coordination Project advisory board, June 8, 2004; also, presented to Cumberland County (ME) Violence Intervention Partnership advisory board, September 8, 2004

To BIP, or not to BIP—That is the question

	page
Research results I: True experiments	4
Critique of the true experiments	
Palmer experiment	
Davis experiment	
Dunford experiment	
Feder experiment	
Research results II: BIP completers vs. BIP dropouts	7
Completers reoffend less than dropouts do	
Can BIP attendance be improved?	
Judicial monitoring & sanctioning	
Motivational enhancement by BIP providers	
Will improving attendance reduce reoffending?	
Conclusion: Reasons to BIP	10
1. <u>REDUCED REOFFENDING: When more offenders complete batterer education programs there will probably be fewer victims of domestic violence.</u>	
2. <u>IMPROVED RISK MANAGEMENT: Monitoring BIP attendance will improve risk management of DV offenders.</u>	
3. <u>PROOF OF CHANGE: By completing BIPs, offenders can demonstrate their commitment to change.</u>	
References	12
Data tables	15
Endnotes	18

To BIP, or not to BIP--That is the question

The most obvious reason To BIP—To assign Domestic Violence offenders to Batterer Intervention Programs--is because the programs work. That is, because men who complete a BIP will reoffend less often than men who don't. But do the programs work? What does research tell us? Let's begin with some conflicting answers to this question:

1. The director of grants at a regional foundation recently stated:

“I've been advised not to provide any further funding to batterer programs because they don't work. I've been told that program evaluations show “no effect” over just putting a man on probation.” (Cited in Gondolf, 2002, p. 28)

2. But in 2000 the author of the book Changing Violent Men concluded:

“The men who completed the abuser programs were significantly more likely to reduce these [violent] acts than men sanctioned in other ways. This strongly suggests that abuser programs are much more successful than other forms of criminal justice interventions.” (Dobash, Dobash, Cavanagh, & Lewis, 2000, pp. 118, 123)

3. In the same year, though, the author of an experimental study of 861 Navy men in San Diego summarized his findings:

“All of the assessments made...point to the same conclusion: The batterer interventions of the cognitive-behavioral model failed to produce meaningful changes in the behavior they were designed to impact.” (Dunford, 2000, p. 475)

4. But in 2002, the author of a multi-site study of more than 800 batterers reported, in the book Batterer Intervention Systems:

“We found a consistent and substantial program effect using three different analyses....Moreover, the moderate effect size was higher than in most previous batterer program evaluations, especially the recent experimental evaluations.” (Gondolf, 2002, p. 144)

5. Last summer, however, a National Institute of Justice report concluded:

“The methodological limitations of virtually all these evaluations make it impossible to say how effective BIPs are.” (Jackson et al., 2003, p. 1)

6. And this year a review of all studies that include a control condition concluded:

“In general, the effect size due to group battering intervention on recidivism of domestic violence is in the ‘small’ range.” (Babcock, Green, & Robie, 2004, p. 1043)

So...research on BIPs shows: That BIPs are effective, that BIPs are not effective, and that it is impossible to say whether BIPs are effective or not!

Now let's examine some of this research for ourselves. First, what kind of study could tell us whether batterer education is effective?

Part I: True experiments

The gold standard for determining whether any kind of intervention works is called a randomized experiment. With this methodology a sample of people is drawn from a known population and randomly assigned to either an experimental group that gets a treatment, like BIPs, or to a control group, that doesn't get the treatment.

If the BIP group later recidivates less, we can say that exposure to the BIP caused the difference—because the random assignment should ensure that there are no other consistent differences between the groups. And we can generalize this causal inference to the larger population from which the sample was drawn—but trying to generalize beyond that sample can be problematic.

Four such randomized experiments have tried to test the effectiveness of BIPs. I have briefly summarized them in Table 1.

Reading across the columns, from left to right, Table 1 lists:

1. The author, date & location of each experiment.
2. The Experimental group: These are offenders who were (randomly) assigned to a BIP (usually along with probation)
3. The Control (no BIP) group. These are offenders who were (randomly) assigned to receive only Probation or some other non-BIP experience.
4. The Type of data examined to see if assignment to a BIP caused a difference in reoffending.
5. The last column on the right tells us whether, in each case, the data showed that the experimental group recidivated less than the control group did.

Results & critiques of the true experiments

There is a “Yes” in the first two rows of Table 1--evidence that Bips worked there, and a “No” in the remaining six rows --indicating no evidence that Bips worked. So the results of these experiments seem mixed, with the preponderance of the evidence not showing that Bips are effective.

When we take a closer look at these experiments, however, we will see that methodological problems and offender sampling limitations prevent us from drawing any meaningful conclusions--YES or NO--about BIPs effectiveness from their results.

Let's consider these experiments row by row: Due to time limitations I'm going to identify only one problem with each of these experiments. (The reader who wishes a more detailed critique is referred to the endnotes in this section.)

The Palmer experiment

Looking at the far right column entry for the Palmer experiment, we see that the BIP group did reoffend less than the control group did. The difference was statistically significant--and sizeable: In a 1-2 yr follow-up, offenders in the control group were three times as likely to reoffend as offenders who were assigned to the BIP.ⁱⁱ

Critique: Methodologically, this experiment was conducted pretty well. The major problem is the very small size of the sample—a total of only 59 offenders. It is unlikely that results from such a small sample would be representative enough to support any broad generalizations about BIPs effectiveness.

The Davis experiment

Police data for the 26 week BIP group also indicate that the education was effective in the Davis experiment. And, as in Palmer's study, this effect was statistically significant and sizeable: One year later 26% of the men in the control group had reoffended compared to only 10% in the BIP group.ⁱⁱⁱ But none of the other Brooklyn comparisons found any support for the effectiveness of a BIP.

Critique: In the Brooklyn experiment, offenders who failed to attend the BIP, as required, were rarely sanctioned for their noncompliance. (As the authors explained it, by the time a pattern of nonattendance had been noted and the information was passed along from the provider to probation to the prosecutors, the defendant was often nearing the end of his probation and the D.A.s didn't bother to pursue the case.) In contrast, however, when men in the community service (control) group didn't show up for work a warrant was issued for their arrest! Thus at least some offenders in the experimental group were essentially learning that they could violate court orders with impunity while offenders in the control group were learning just the opposite lesson. Other things being equal, then, we might expect these lessons to cause the control group to reoffend less often than the BIP group.^{iv} In any event, we obviously don't have a level playing field for comparing the BIP groups to the control group—a serious violation of the requirements of a true experiment.

The Dunford experiment

Dunford's study has fewer methodological problems than the Brooklyn experiment, and Table 1 shows that he found no evidence at all that a BIP was effective.

Critique: The problem here is that serious questions must be raised about whether we can generalize from Dunford's sample of offenders to offenders in any criminal justice

jurisdictions in this country. Let me profile the offenders in his study and you can tell me whether you recognize these men:

- 1.--Few if any had criminal histories--especially felonies (since the Navy usually attempts to screen them out at enlistment).
- 2.--None had substance abuse problems (they had been screened out or treated before being admitted to the experiment).
- 3.--None had identifiable mental health issues, including “pathological jealousy”.
- 4.--All of them were employed (by the U.S. Navy).
- 5.--All of them lived in a structured community provided by their employer.
- 6 --100% were married (this was set up as a couples study).
- 7.--None had divorce proceedings in progress.

Does anybody recognize this group of offenders? In research conducted in many different jurisdictions across the country that I have read for this presentation I certainly never encountered a sample of offenders that matched more than one of the characteristics in this profile.

So I think it’s safe to say that these offenders do not remotely resemble the usual suspects. Therefore, we can’t really draw any conclusions from the Dunford experiment about the effectiveness of batterer education programs in criminal justice jurisdictions in this country.^v

The Feder experiment

The last experiment listed in Table 1 also found no difference in reoffending between the group that received the batterer education and the group that did not. In an improvement over Dunford’s study the offenders in Feder’s experiment do appear to at least resemble offenders in many urban criminal justice jurisdictions in the U.S.

Critique: The men who got randomly assigned to the control group were not allowed to enroll in a BIP program, so many criminal justice players in Broward Co. saw the random assignment, the experiment, and the researchers themselves, as compromising victim safety. As a consequence, victim advocates, probation, and prosecutors alike were openly hostile to the researchers. This compromised the experiment:

Feder & Forde (2000, p. 125) state: “..We had to deal with actions taken by various courthouse personnel aimed at thwarting the study. So, for instance, we would begin speaking with a victim about the interview when one of the assistant prosecutors would come over to the woman and explain that we were the reason that the judge was not placing her partner into counseling. That it was our study that was responsible for placing her in danger.” As a result of what Feder called this “hostile environment” it’s not surprising the researchers ended up with only about a 25% response rate from their victims—far too low to draw meaningful comparisons between the experimental and control groups in their study.^{vi}

So despite the considerable lengths all these investigators went to in meticulously planning these experiments (Dunford spent four years just selecting the sample for his study!), they all ran into trouble either in executing their designs or generalizing from their results.

Furthermore, in my opinion it's going to be a very long time before we get usable results from any true experiments in this field for the following reason: When we can control events and players well enough to do a methodologically sound experiment we probably aren't working in the real world, but when we are working in the real world then we probably can't control events and players well enough to do a methodologically sound experiment. This Catch-22 has implications for victim safety, which I'll touch on in my conclusion.

It is my impression that it is these experiments that people most commonly cite when they conclude that "BIPs don't work". In fact, because these experiments are all fatally flawed they cannot provide evidence, one way or the other, about BIPs effectiveness. But results from a different methodology do strongly suggest that BIPs are effective. I turn now to these results.

Part II: BIP Completers vs. BIP Dropouts

This research compares recidivism of men who complete (or nearly complete) a batterers' program to offenders who drop out of (or never show up at) the program. The logic of this design is that if BIPs work then completers, who get more batterer education than dropouts, ought to reoffend less than dropouts do.^{vii} First, let's see if that's true.

In Table 2 I have listed every study I could find that appeared in the last decade that compared reoffense rates of BIP completers to reoffense rates of BIP dropouts. Reading from left to right: the first column lists the study, date, and location; then the type of data (Police or Victim interviews, or sometimes both); then, in the next two columns, the percentage of dropouts reoffending (after some specified time period) and the percentage of completers reoffending (in the same time period). In the last column, a "Yes" indicates that completers did reoffend less than dropouts did.

Table 2 assembles a large and extremely diverse set of data. There are results from many parts of the country: East, West, South, and Midwest. A variety of offender samples are also represented here: Some are predominantly white, some predominantly black, one is mostly Hispanic. In some samples most offenders have a criminal history, only a minority of men in other samples has previously offended. In some samples most of the offenders were charged with DV felonies, in other samples they were nearly all misdemeanor DV charges. In one sample nearly all the offenders were employed, in another sample only half of them were. Police data as well as reports from victims are represented here. Furthermore, the sample is huge: Over 6,000 offenders were observed in these studies. And a dozen different investigators conducted the research.

Completers reoffend less often than dropouts do

Yet despite this tremendous diversity, one thing doesn't vary: In every single case completers reoffend less often than dropouts do. I have read more than 300 studies in the field of domestic violence and this is the most consistent set of data I have ever seen. And it is not a small effect. Averaging over all the studies assembled in Table 2, dropouts are more than twice as likely to reoffend as completers are.

This completion effect is large, but a BIP is not a magic bullet. Roughly 20% of the BIP completers represented in Table 2 did reoffend. Nevertheless, completers reoffend much less often than dropouts do—and statistically controlling for other observed differences between completers and dropouts (e.g., in employment, criminal history) does not eliminate the difference in reoffending^{viii}. This evidence raises the distinct possibility that a strategy of moving men from the dropout column to the completion column will reduce reoffending overall. And mandating even more men to a BIP, and getting them to complete it should help even more. But can we really get more men to complete a BIP?

Can BIP attendance be improved?

Well, there is certainly plenty of room for improvement in attendance: Only a little more than half of the offenders reported in Table 1 and Table 2 actually completed the programs they were mandated to attend and this is very much in line with other published surveys (Daly & Pelowski, 2000; Pirog-Good & Stets, 1986). Even so, maybe everybody in this recalcitrant population who is going to complete a BIP is already completing it. But research shows otherwise. It turns out that it is actually not very difficult—or expensive—to substantially improve BIP completion rates.

Judicial monitoring and sanctioning

For example, judicial monitoring and sanctioning can improve BIP completion rates: As some of you probably already know, a study conducted at the Pittsburgh DV court found that completion rates shot up from one-half to two-thirds soon after a policy of judicial monitoring coupled with swift sanctions for non-compliance was instituted (Gondolf, 2000). And this is our own anecdotal experience here in Maine in the Portland and York DV case coordination projects.

Motivational enhancement by BIP providers

Researchers in Howard Co., MD (the Taft study in Table 2) took a different approach to improving BIP attendance. They adopted “motivational enhancement” techniques that have brought about big increases in attendance at substance abuse programs.

In their study a BIP leader did immediate and personal follow-up with clients who missed sessions. These follow-ups included handwritten notes, phone calls, expressions of concern about the client not being there, telling him that others in the group had missed him, reminding him of the possible penalties for not completing, etc.

Regardless of what you think of this approach, it did seem to work: Even though the dropout rates in this jurisdiction were quite low to begin with, instituting this motivational enhancement technique cut the existing dropout rates in half (from 30% to 15%). So this approach did get more offenders to complete the program. And the completers were still much less likely to reoffend than the dropouts were (as you can see in Table 2).

So research shows that we can increase BIP completion rates. And given the very strong connection (documented in Table 2) between completing a BIP and being less likely to reoffend, it's at least a good bet that getting more men to complete the programs will reduce the overall tendency to reoffend.

Will improving attendance reduce reoffending?

Can I cite any research showing that this will happen? Yes.

A study published late last year, in the journal Criminology & Public Policy, compared recidivism before and after a DV court went operational in Lexington Co., SC (Gover, MacDonald, & Alpert, 2003). This report included some very nice controls that make it much stronger than the usual before and after study.

This DV court, which was part of a coordinated community response team, handled all non-felony DV battery cases in the county and placed a strong emphasis on mandating offenders to a 26 wk BIP, combined with strict weekly follow-ups on the offenders' progress, and it included sanctions (imposing a suspended jail sentence) if they failed to comply.

The researchers compared cases that were processed before the DV court started to cases that were processed through the DV court. They didn't report BIP completion rates, but based on the Pittsburgh study and on our experience here, I think we can pretty safely assume that more men completed the BIP program after the DV court was in place than before it started. The researchers did compare these offenders--on demographics, criminal history, etc., but the only reliable difference between them was that offenders who were processed through the DV court had significantly lower DV recidivism during an 18 month post-arrest window. This drop in recidivism did not just reflect a drop in DV in that jurisdiction, because DV arrests, overall, actually increased during this period.^{ix}

Here is what that drop in reoffending meant to victims in Lexington County: Over the three year period when the court was supported by a VAWA grant, they processed 2500 cases. Based on the before and after recidivism rates they reported I calculated that during this period more than 200 women avoided the assaults and in some cases serious injuries that they would have suffered without the DV court.^x

And that's in only one county in one state. Even a much smaller effect, nationally, could benefit many more victims. For example, using Bureau of Justice Statistics estimates of nearly a million DV crime victims annually, Babcock et al. (2004) calculated that even a

5% drop in reoffending would mean that 42,000 women would avoid being criminally abused every year.

Reasons to BIP

In conclusion, although experiments on BIPs effectiveness are inconclusive, the research I have reviewed in the second part of this presentation provides two compelling reasons for making offenders complete a BIP.

1. REDUCED REOFFENDING: When more offenders complete batterer education programs there will probably be fewer victims of domestic violence.

All the non-experimental research conducted in this decade shows that offenders who complete a batterers' program are less likely to reoffend than are offenders who drop out. Controlling for all other differences between completers and dropouts that researchers have been able to think of so far does not make this effect go away. The obvious implication of this research is that if more offenders completed batterer education there would be fewer victims of domestic violence.

Although this kind of research does not definitively prove that BIPs work, victims are at risk right now. Many women will be punched in the face, thrown down the stairs, kicked in the stomach when they are pregnant--and even beaten to death, as Lisa Deprez was while I was preparing this report—if we postpone action until we have definitive proof.

Because there is good presumptive evidence that BIPs work, I believe that the Precautionary Principle (Raffensperger & Tickner, 1999), borrowed from environmental law, should guide our actions. This principle states:

“When an activity raises threats of harm to human health or the environment, [as DV surely does] precautionary measures should be taken even if some cause and effect relationships are not fully established scientifically.”^{xi}

2. IMPROVED RISK MANAGEMENT: Monitoring BIP attendance will improve risk management of DV offenders.

Quite apart from the issue of whether BIPs truly reduce reoffending is the fact that dropping out of a BIP is a red flag for reoffending. In fact, dropping out predicts reoffending more consistently than any other risk factor that research has yet identified.^{xii} The only way we can obtain this particular information is to assign an offender to a BIP and monitor his attendance. Then, if this red flag is raised, swift criminal justice responses such as heightened scrutiny and incarceration may prevent reoffending. At the same time victims can be alerted so that they can review their safety planning in light of the increased danger.

3. PROOF OF CHANGE: Completing a BIP demonstrates an offender's commitment to change.

Now I want to go beyond these two evidence-based reasons and offer you a third, somewhat different justification for mandating BIPs. A senior member of the Maine Judiciary recently remarked: “We need to make DV offenders accountable on many levels, and assigning them to BIPs is something we can do to make them prove that they have changed.” Mandating men to BIPs provides them with “An Opportunity for Change”, to borrow the name of a Cumberland County program. A man who completes a BIP demonstrates a willingness to change. He can be encouraged and rewarded for his commitment and perhaps this will put him on track for a violence-free life. If mandating men to BIPs can even sometimes achieve this result, then it is an option much to be recommended.

References

- Baba, Y., Galaka, S., Turk-Bicakci, L., & Asquith, D. (1999). Variables associated with success or failure in a court-ordered domestic violence treatment program. Downloaded Oct. 15, 2003 from: <http://www.growing.com/nonviolent/research/dvprog>.
- Babcock, J. C., Green, C. E., & Robie, C. (2004). Does batterers' treatment work? A meta-analytic review of domestic violence treatment. *Clinical Psychology Review*, *23*, 1023-1053.
- Babcock, J. C. & Steiner, R. (1999). The relationship between treatment, incarceration, and recidivism of battering: A program evaluation of Seattle's coordinated community response to domestic violence. *Journal of Family Psychology*, *13*, 46-59.
- Coulter, M., Byers, R., Menezes, L., & Jayakumar, A. (2001). An evaluation of the Hillsborough County domestic violence intervention program of the 13th judicial court. Technical Report. Contact: mcoulter@hsc.usf.edu
- Daly, J. E., & Pelowski, S. (2000). Predictors of dropout among men who batter: A review of studies with implications for research and practice. *Violence and Victims*, *15*, 137-160.
- Davis, R. C., Taylor, B. G., & Maxwell, D. D. (2000). Does batterer treatment reduce violence? A randomized experiment in Brooklyn. Final report for National Institute of Justice, grant 94-IJ-CX-0047. Washington, DC: National Institute of Justice. NCJRS. NCJ 180772.
- Dobash, R. E., Dobash, R. P., Cavanagh, K., & Lewis, R. (2000). *Changing violent men*. Thousand Oaks, CA: Sage.
- Dunford, F. W. (2000). The San Diego Navy experiment: An assessment of interventions for men who assault their wives. *Journal of Consulting and Clinical Psychology*, *68*, 468-476.
- Feder, L., & Forde, D. R. (2000). A test of the efficacy of court-mandated counseling for domestic violence offenders: The Broward experiment. Final report for National Institute of Justice, grant number 96-WT-NX-0008, Washington, DC: National Institute of Justice. NCJRS. NCJ 184752.
- Gondolf, E. W. (1997). Patterns of reassault in batterer programs. *Violence and Victims*, *12*, 373-387.
- Gondolf, E. W. (2000). Mandatory court review and batterer program compliance. *Journal of Interpersonal Violence*, *15*, 428-437.

Gondolf, E. W. (2001). Limitations of experimental evaluation of batterer programs. Trauma, Violence, & Abuse, 2, 79-88.

Gondolf, E. W. (2002). Batterer intervention systems: Issues, outcomes, recommendations. Thousand Oaks, CA: Sage.

Gordon, J. A. & Moriarty, L. J. (2003). The effects of domestic violence batterer treatment on domestic violence recidivism: The Chesterfield county experience. Criminal Justice and Behavior, 30, 118-134.

Gover, A. R., MacDonald, J. M., & Alpert, G. P. (2003). Combating domestic violence: Findings from an evaluation of a local domestic violence court. Criminology and Public Policy, 3, 109-132.

Jackson, S., Feder, L., Forde, D. R., Davis, R. C., Maxwell, C. D., & Taylor, B. G. (2003, June). Batterer intervention programs: Where do we go from here? NIJ Special Report. Washington, DC: National Institute of Justice. NCJRS. NCJ 195079. May be downloaded from: <http://www.ncjrs.org/pdffiles1/nij/195079.pdf>.

Murphy, C. M., Musser, P. H., & Maton, K. I. (1998). Coordinated community intervention for domestic abusers: Intervention system involvement and criminal recidivism. Journal of Family Violence, 13, 263-284.

National Institute of Justice (2003, September). Do batterer intervention programs work? Two studies. NIJ Research for Practice. Washington, DC: National Institute of Justice. NCJRS. NCJ 200331. May be downloaded from: <http://www.ncjrs.org/pdffiles1/nij/200331.pdf>

Palmer, S. E., Brown, R. A., & Barrera, M. E. (1992). Group treatment program for abusive husbands: Long-term evaluation. American Journal of Orthopsychiatry, 62, 276-283.

Pirog-Good, M. A., & Stets, J. (1986). Programs for abusers: Who drops out and what can be done. Response to the Victimization of Women & Children, 9, 17-19.

Puffett, N. K. & Gavin, C. (2004, April). Predictors of program outcome & recidivism at the Bronx misdemeanor domestic violence court. New York: Center for Court Innovation. Downloaded April 29, 2004 from: http://www.courtinnovation.org/pdf/predictors_bronx_dv.pdf

Raffensperger, C., & Tickner, J. (1999). Protecting public health & the environment: Implementing the precautionary principle. Washington, DC: Island Press.

Rosenbaum, A., Gearan, P. J., & Ondovic, C. (2001). Completion and recidivism among court- and self-referred batterers in a psychoeducational group treatment program:

Implications for intervention and public policy. Journal of Aggression, Maltreatment & Trauma, 5, 199-220.

Shepard, M. F., Falk, D. R., & Elliott, B. A. (2002). Enhancing coordinated community responses to reduce recidivism in cases of domestic violence. Journal of Interpersonal Violence, 17, 551-569.

Taft, C. T., Murphy, C. M., Elliott, J. D., & Morrel, T. M. (2001). Attendance-enhancing procedures in group counseling for domestic abusers. Journal of Counseling Psychology, 48, 51-60.

Table 1: Does batterer education work? The four true (at least in conception) experiments

In a properly conducted true experiment, offenders would be randomly assigned to either a batterers’ education group, all of whom would get batterers’ education, or to a control group who would not receive any batterers’ education. Any resulting differences between the groups (a “Yes” in the last column) could then be attributed to the batterers’ education, since there should be no other consistent differences between the groups that could explain the result. The absence of an effect (a “No” in the last column) would suggest, but could not prove (because the non-existence of an effect cannot be proven), that the batterers’ education did not work.

Unfortunately, practical and ethical considerations can prevent a true experiment from being conducted properly, or sampling limitations may prevent us from generalizing its results. When this occurs, as it did to some degree in all of the studies listed below, cause and effect inferences can no longer be made with confidence, and the absence of a difference does not imply the absence of a treatment effect.

Experiment / location	Experimental group	Control group	Type of data	Did BIP educ. reduce reoffending?
Palmer (1992) / Ontario, CANADA	Probation + 10 wk BIP	Probation only	Police	Yes

Davis (2000) / Brooklyn, NY	40 hrs of BIP (in 8 weeks or 26 weeks)	40 hrs of community service	Police	Yes (for 26 wks group)
			Police	No (for 8 wks group)
			Victim reports	No (both groups)

Dunford (2000) / San Diego, CA	30 wk BIP	Safety planning for victims	Police	No
			Victim reports	No

Feder (2000) / Broward Co. FL	Prob. + 26 wk BIP	Prob. only	Police	No
			Victim reports	No

Notes for Table 1:

a. BIP completion rates per study:

Palmer = 70% Davis = 40% Dunford = 71% Feder = 66%

b. Number of offenders per study:

Palmer = 59 Davis = 376 Dunford = 318 Feder = 404

Table 2 (next page): Does batterer education work? Non-experimental studies that compare dropouts to completers

Table 2 shows that BIP completers are much less likely to reoffend than are BIP dropouts. This means that dropping out of a batterers’ program is a clear risk factor for reoffending. The studies cited in Table 2 conclusively establish this fact. But they do not necessarily prove that batterers’ education works.

Maybe completers do reoffend less often because they are exposed to more batterers’ education than dropouts are. That is, they reoffend less often because batterers’ education works. This explanation seems likely but, because the studies cited in Table 2 are not true experiments, alternative explanations are also possible. For example, some other differences between completers and dropouts may explain the differences in reoffending. Yet some of the studies cited in Table 2 found no discernable differences (e.g., no differences in criminal history, age, employment, substance abuse), between completers and dropouts--but even after these differences were statistically controlled for completers still reoffended less than dropouts did. Thus, observed differences between completers and offenders cannot adequately explain the “Yes” entries in Table 2.

Of course completers and dropouts may differ in unknown ways and it could be these unknown, pre-existing, differences, rather than exposure to different amounts of batterers’ education, that explain the different reoffense rates. Until these unknown differences are documented, however, the best currently available explanation for the differences in reoffense rates documented in Table 2 is that batterers’ education works. Therefore assigning more batterers to Bips and ensuring that they attend seems to be a promising strategy for reducing domestic violence.

Notes for Table 2:

a. Table 2 lists every study reported in the last decade (published as well as unpublished)--that could be located after a diligent search--that compared the reoffense rate for BIP completers to the reoffense rate for BIP dropouts.

b. For most of the studies reoffending refers to DV reoffending, but a few studies reported any new offenses.

c. “Completers” was defined by the authors of each study and usually meant attending most, but not all, BIP sessions. Completion rates ranged from 16% (Murphy, 1998) to 85% (Taft, 2001).

d. BIPs were usually Duluth or Cognitive Behavioral or hybrid. (The few Anger Management programs that were located are not included in this table, but all of them found the same effect that is reported here.)

e. Average reoffense rates:	<u>Dropouts</u>	<u>Completers</u>
By Police report:	32%	12%
By Victim report:	51%	33%

Table 2: Reoffending for BIP dropouts vs. BIP completers

Study location	Type of data	% of dropouts who reoffended	% of completers who reoffended	Did BIP completers reoffend less than dropouts?
Murphy (1998) Baltimore, MD	Police	16%	0%	Yes
Baba (1999) Santa Clara Co., CA	Police	8%	1%	Yes
Babcock (1999) Seattle, WA	Police	23%	8%	Yes
Dunford (2000) San Diego, CA	Police / Victim	(% reoffending not reported)		Yes (but “very small” effect)
Feder (2000) Broward Co., FL	Police	30%	13%	Yes
Coulter (2001) Hillsborough Co., FL	Police	12%	6%	Yes
Rosenbaum (2001) Central MA	Police	14%	3%	Yes
Taft (2001) Howard Co., MD	Police Victim	54% 33%	10% 15%	Yes Yes
Gondolf (1997, 2002) Dallas, TX	Police Victim	19% 58%	12% 33%	Yes Yes
Denver, CO	Police Victim	51% 55%	26% 35%	Yes Yes
Houston, TX	Victim	59%	35%	Yes
Pittsburgh, PA	Police Victim	41% 50%	17% 40%	Yes Yes
Shepard (2002) Duluth, MN	Police	51%	40%	Yes
Gordon (2003) Chesterfield Co., VA	Police	(% reoffending not reported)		Yes
Puffett (2004) Bronx, NY				
BIP only group	Police	47%	14%	Yes
BIP & S.A. group	Police	48%	9%	Yes

Endnotes

ⁱ Robert Moyer, Emeritus Professor of Psychology, Bates College, is a member of the Maine Commission on Domestic and Sexual Abuse and the advisory boards of: the Cumberland County Violence Intervention Partnership, the Portland DV case coordination project, the York/Springvale DV case coordination project, and the Bail Commissioners DV Training Project. Please send comments on this ms. to: rmoyer@bates.edu

ⁱⁱ Because there was substantial attrition in the Palmer study (i.e., many offenders who were mandated to complete the batterer program did not comply), comparisons of the experimental to the control group very likely underestimate the size of the true effect. This is because the logic of experimental design requires that the BIP dropouts be treated as if they completed the program. Naturally, this should diminish the observed impact of the BIP. (It's like measuring how much a daily dose of aspirin reduces heart attacks when participants in your study actually take their aspirin only 4 days a week.). This is a problem for every experiment listed in Table 1, and it means that, in part, all the data analyses are asking how much people benefited from the batterers' education they didn't get!

ⁱⁱⁱ It is puzzling that a recent National Institute of Justice assessment (2003, September, p. 1) calls this improvement—a nearly 2/3 reduction in recidivism-- “only minor”. If Congress proposed cutting the DOJ budget by nearly 2/3 I do not think the folks at the NIJ would regard the reductions as “only minor”.

^{iv} Another problem with the Davis experiment was caused by judicial overrides: 14% of the offenders who were supposed to be assigned to the control group instead were assigned by judges to the BIPs group. As the authors note: “Substantial concessions had to be made to court officials to gain their cooperation.” But the data analysis treated these overrides as if they had had no BIPs exposure. This means that if the BIPs truly did have an effect, it would be harder to detect it, because some of the Controls who had been misassigned to a BIP would, as a result, be reoffending less often too.

^v Some may argue that the fact that batterer education didn't work for Dunford's offender sample, which had such a high “stake in conformity”, is conclusive evidence that it also certainly won't work for a more representative sample of offenders (who have much less of a stake in conformity). But this argument cuts both ways: This high stake in conformity may have so effectively reduced reoffending in the control group, as well as in the experimental group, that it became difficult to see any differences between the groups. Consistent with this response, Dunford did document large reductions in DV reoffending in all groups (for simplicity I have reported only the results of his two crucial groups here) after the interventions; it's just that the magnitude of these reductions didn't vary by

group. But only repeating his study on a more representative sample of offenders can resolve this debate.

^{vi} But the Broward Co. experiment also failed to find an effect based on Police data. Yet they must have underestimated any effect of BIPs that could have been present. First, because half of the observation period for measuring reoffending had ended before any offender could have completed the 6 month BIP. So, many of the recorded offenses most likely occurred before many offenders got much of a dose of BIPs. This problem, coupled with their 29% attrition rate, would be expected to dilute potential effects of batterer education. Another problem is that some of the police data they report are arrests that were based on probation violations; but the BIPs group could (and often was) violated for failure to attend BIPs sessions—something the Control group could not be violated for. The authors do some analyses to try to blunt the impact of this problem but I don't believe they entirely succeed. Gondolf (2001, p. 83) also calls attention to this difficulty.

^{vii} To say that men who get more batterer education should reoffend less than men who get less does not mean that we must also expect longer programs to be more effective than shorter ones: Driver re-education courses might be very effective at reducing accidents, but a 12-week course might be no more effective than a 6-week course. To date, research has not shown that longer BIP courses are more effective than shorter ones, though clean comparisons are difficult because the programs, and the criminal justice contexts in which they are offered, usually differ in many ways.

^{viii} Eight of the studies reported in Table 2 did measure a variety of offender characteristics that could be related to reoffending—employment, criminal history, etc.—but in only one study (Feder & Forde, 2000) were they able to eliminate the BIP completion effect when they statistically controlled for these differences. On balance, then, the data are quite consistent with the claim that it is the BIP—and not some other difference between completers and dropouts--that is responsible for the big difference in reoffending.

^{ix} Furthermore, the drop in reoffending was specific to DV recidivism. That is, men processed through the DV court were just as likely to be subsequently arrested for non-dv assault as were the men who had not been through the dv court. So the mandated BIP with sanctions (and probably the Coordinated Community Response it was part of) had a focused effect on DV recidivism.

^xI made this calculation as follows: In the 3 years the DV court was supported by a VAWA grant they processed about 2500 cases. Before the DV court was formed, we may estimate (since the sampled cases were a random sample of all cases processed) that 450 of those offenders (18% of 2500 pre-DV court cases) would have recidivated, but only 250 would have done so after the court was in operation (10% of 2500 post-DV court cases). So over this 3-year period we might expect that at least 200 women avoided

abuse. And this calculation pertains only to cases that were severe enough to make it to court.

^{xi} I am grateful to Cathy Lee for bringing the Precautionary Principle to my attention.

^{xii} Although dropping out appears to be the most consistent risk factor for reoffending, other risk factors may, in some studies, be larger. For example, Gondolf (2002) found that men in BIPs who were “drunk every night” were about 16 times as likely to re-assault their partners as were men in the programs who seldom or never drank—none of the BIP completion effects listed in Table 2 even approach that magnitude. Yet Puffett & Gavin (2004) found that substance abuse did not significantly predict recidivism in their study. Nevertheless, both these studies did find that dropping out of a BIP significantly predicted recidivism. Some of the inconsistent findings from risk factors such as criminal history, employment, and substance abuse, that often do show substantial effects may reflect differences in how these risk factors are measured, including whether the measures are contemporaneously made, as well as differences in offender samples and criminal justice context. The BIP completion effect that I have documented in Table 2, however, seems to be so robust that it transcends all these influences.