Policing Domestic Violence 1967–2017

Lawrence W. Sherman
University of Maryland—College Park and University of Cambridge

Abstract

The U.S. President’s Commission on Law Enforcement and Administration of Justice under President Johnson in 1967 called for a program of research that would support evidence-based tracking, targeting, and testing of policing domestic “disputes.” During the past 50 years, the amount of research on domestic violence has grown. The findings from targeting studies reveal a steep pyramid from many low-harm cases at the base rising to a tiny fraction that are lethal. The findings from testing studies reveal that arrests reduce reoffending among employed persons in the short run while increasing recidivism among unemployed suspects. Arrests in one experiment increased all-cause mortality of victims over 23 years. The findings from tracking research reveal wide variance across police agencies in implementing mandatory arrest. Since the Crime Commission, however, mandatory arrest laws were also enacted in 28 states in the 1980s, with uneven enforcement for misdemeanor domestic abuse. New research is needed to test the effectiveness of other domestic violence interventions, including restorative justice, as well as follow-up studies of previous research sites.

Keywords
policing, domestic abuse, mandatory arrest, second responders, victim impact

The U.S. President’s Commission on Law Enforcement and Administration of Justice (commonly known as the President’s Crime Commission or just “the Commission”) wrote very little about policing domestic violence in its 1967 report (defined there and in this article as “intimate partner violence”). That fact must be considered in the context of powerlessness of women at that time. The Commission...
offered no substantive discussion of crimes against women, except to note they were a major
category of homicides. Its entire, if scanty, approach to domestic abuse emphasized the
problem it created for police, not for women.

The best evidence for this conclusion is the complete absence of any reference to
the substantive problem of domestic violence in the index of either *The Challenge of
Crime in a Free Society*, the main report of the Commission (1967b), or its *Task Force
reveals no entries under the following terms: police, domestic, family, women, disputes,
disturbances, homicide, murder, restraining order, or child. Checking the same terms in *The
Challenge of Crime* reveals nothing on the harm to women, but on two pages of the report
(pp. 92 and 104), “Police intervention, situations requiring, matrimonial disputes” was
discussed.

Yet what little the Commission did say in the report (1967b) about the issue was
extraordinarily helpful. The content of the report has guided (or at least has predicted)
important directions that research has taken since then. In particular, the Commission
identified the issues of whether arrest was an appropriate response, as well as how to predict
domestic homicide. The full text of what the Commission report said on pp. 92 and
104 is provided in the appendix. What follows are selected direct quotations from the
report:

A common kind of situation that illustrates the complexity, delicacy and frustra-
tion of much police work is the matrimonial dispute, which police experts
estimate consumes as much time as any other single kind of situation. . . . Yet
the capacity of the police to deal effectively with such a highly personal matter
as conjugal disharmony is, to say the least, limited. Arresting one party or both
is unlikely to result in either a prosecution or a reconciliation. Removing one
of the parties from the scene . . . may create temporary peace, but it scarcely
solves the problem. . . . [M]ediating the difficulty of enraged husbands and
wives ad hoc is an activity for which few policemen—or people in any other
profession—are qualified by temperament or by training. (p. 92)

Fourteen years later, these three options—arrest, removal, or mediation—were the
focus of a 1981–82 National Institute of Justice (NIJ) grant for the first randomized field
experiment in police domestic violence (Sherman and Berk, 1984a, 1984b). Two other lines
of testing police responses to domestic abuse were anticipated by the Commission when it
said that police should “compile statistics on the typical effects of having one of the parties
swear out a complaint against the other, to become familiar with the social-service agencies,
if any, to which troubled families can be referred” (1967b: 104).

Those recommendations by the Commission (1967b) anticipated the research by
Dunford (1990) and by the authors reviewed in Davis, Weisburd, and Taylor (2008). Yet
the Commission’s (1967b: 104) report went on to discuss the equally important issues
of tracking and targeting, as well as of testing, noting that “[t]he police should seek to accumulate information about families that cause repeated disturbances, to discover whether certain kinds of disturbances are more likely than others to lead to serious assaults or to homicides.” Perhaps without any of the police researchers concerned ever realizing it (including this author), these short statements pinpointed the key directions of research for the next 50 years. The Commission asked what now seem to have been the right questions, even if it asked them for the wrong (or not primary) reasons of reducing harm to victims.

**Major Research Findings During the Last Half-Century About Policing Domestic Violence**

The major research findings fit clearly into the three “Ts” of evidence-based policing: Targeting, Testing, and Tracking (Sherman, 2013). As some of the best work in this area has been done in other countries, it is included here. The purpose of the inclusion is to point out the many similarities in findings across countries, as well as to invite further replications and comparisons to advance the external validity of conclusions from these lines of research.

**Targeting**

The findings from large-sample studies of domestic abuse reported to police consistently reveal a steep pyramid, with massive numbers of low-harm cases at the base and declining numbers of more serious cases rising to a tiny fraction that are lethal. Few studies reveal any pattern of escalating seriousness of cases reported to police because most cases have no repeat calls, and most study results show that there is no escalation in severity of injury when repeat cases do occur. In contrast, in the minority of cases with repeated reports, the findings do show increasing frequency with each new event, but that does not predict any rising risk of murder. Although most domestic murders occur without prior police contact with the offender–victim dyad, the best predictor of domestic homicide now seems to be prior suicide attempts by the killer.

**Crime harm severity distribution.** The pyramidal distribution is more precisely presented in a pie chart of incident-based harm levels (Sherman, Neyroud, and Neyroud, 2016) across approximately 141,000 incidents reported to Thames Valley Police in 2010–2015 (Barnham, Barnes, and Sherman, 2017). Less than 1% of reported incidents could have been punished, at conviction, with more than a 30-day prison sentence. Similar distributions are found in Bland and Ariel (2015); Kerr, White, and Strang (2017); and Sherman, Bland, House, and Strang (2016).

**Escalation in severity?** Bland and Ariel (2015) examined all 36,000+ police callouts to family violence incidents in Suffolk, U.K., from 2009 to 2014. Using the Cambridge Crime Harm Index (Sherman, 2013; Sherman, Neyroud, et al., 2016, as the measure of harm severity, Bland and Ariel found:
[N]o escalation in the majority of cases; 76% of all unique victim and offender units (dyads) had zero repeat calls. Among the cohort of 727 dyads who called police 5 or more times, there was no evidence for statistically significant escalating harm severity, but some evidence of increasing frequency. Less than 2% of dyads accounted for 80% of all domestic abuse harm, but in over half of these highest harm dyads, there had been no prior contact with police regarding domestic abuse. (p. 30)

Kerr et al. (2017) also found no general escalation in severity, but they did find the only case to date of any escalation based on reports to police: among Aboriginal repeat family violence offenders in the Northern Territory of Australia. Even though most Aboriginal offenders had no repeat domestic abuse within a multiyear follow-up, those who did—many of whom were female offenders with male victims—showed a pattern of escalating severity.

Barnham et al. (2017), like Kerr et al. (2017), improved on the Suffolk study by using a standard-length follow-up observation period for each case. Barnham et al. used a standard 731-day observation period for all cases to analyze 52,000 unique intimate partner violence cases from the initial report to Thames Valley Police within the study period. Like Bland and Ariel (2015), they found that more than three quarters of the offenders had no repeat incidents. Barnham et al. found evidence of escalating severity with repeat incidents, with the average sentence length recommended for the reported offenses never rising to 35 days.

Barnham et al. (2017) also found increasing frequency with a rising total number of incidents by the same domestic abuser, measured as declining time between incidents as the total number of incidents increases: from 164 days between first and second to 23 days between 19th and 20th. This pattern has misled many practitioners to confuse severity and frequency, especially in the rare case when serious harm does come after a long and frequent history of low-harm events.

Predicting domestic homicide. Research findings from around the world show that most lethal (and near-lethal) domestic violence occurs without prior events known to police. Sherman, Schmidt, Rogan, and DeRiso (1991) found that in 1987–89, only 1 in 33 couples with a domestic homicide had a record of a report of domestic abuse among 15,527 such reports made to police within that time period (see also Sherman, Schmidt, and Rogan, 1992: 236–237). Sherman and Strang (1996) found a similar pattern with 41,245 nonfatal domestic callouts to police in Victoria (including Melbourne) in the early 1990s, when 74 domestic murders occurred—of which only 3% had a domestic record. Replications of these analyses have resulted in similar findings in Thames Valley, U.K. (Thornton, 2017); Dorset, U.K. (Chalkley and Strang, 2017); all of England and Wales (Bridger, Strang, Parkinson, and Sherman, 2017); Leicestershire, U.K. (Button, Angel, and Sherman, 2017); and Western Australia (Sherman, Bland, et al., 2016).
**Suicide attempts as a predictor of homicide.** In the absence of good prediction from prior police data, the most important finding from this line of epidemiological research has been the prevalence of suicide attempts, threats, or ideation known to have been associated with the offender. Thornton (2017) discovered this first in Thames Valley, but she was limited by imprecise records about when the suicidal risk was reported (before or after attempted domestic murder). Bridger et al. (2017) reported that 40% of males who completed domestic murders had been known to someone prior to the murder, but not generally to the police, as suicidal. Button et al. (2017) found that in 1997–2015, arrest records on 158,379 offenders showed that those who reported suicidal risk before a domestic homicide were three times more likely to be found among 620 lethal domestic offenders than were previously arrested persons who had not been reported as suicidal during a prior arrest. The potential for developing early warning systems for domestic homicide based on suicidal risks may be great, but it cannot easily be fulfilled without a solution to privacy issues surrounding mental health records. The “power few” (Sherman, 2007a) in this pattern is so tiny that both better data and methods are required to improve the accuracy of prediction.

**Random forests forecasting.** A promising advance in targeting methods for prevention of high-harm domestic abuse is the development of “big data” forecasting using “random forests” algorithms of the kind that predict murder and other serious crime (Berk, Sherman, Barnes, Kurtz, and Ahlman, 2009). Far more accurate than linear regression, these models promise fewer false positives, as well as fewer false negatives. The first police forecasting model (to our knowledge) using random forests for serious interpersonal violence (IPV) has recently been developed for the Hampshire Police by Dr. Geoffrey Barnes of the Western Australian Police (personal communication, 2017). If police can reduce the very large false positive rate of current predictions of domestic homicide, then they can help to invest more homicide prevention resources among many fewer people (Sherman, 2007b). And because most domestic homicide cases have had no prior police contacts, the use of such models may help persuade mental health authorities and families to share more predictive information.

**Testing**

**Family crisis specialist training.** The testing of police responses to domestic abuse apparently began with the work of clinical psychologist Morton Bard (1970, 1973; Bard, Zacker, and Rutter, 1972) with the New York City Police. Under U.S. Department of Justice funding, he designed and implemented a specialist uniformed patrol team called the Family Crisis Intervention Unit (FCIU), which was designed to sit down and talk out problems with couples whose “conflict” had generated a call to police (often by neighbors or relatives). When I observed the unit answering calls in 1971, FCIU officers separated the parties and sat down with them to ask for a detailed account of what was wrong—with the argument, the day, or the entire relationship. The discussions ended with referrals of the couple to social service agencies, on a voluntary basis. This idea spread around the
country, but it was never rigorously evaluated with any IPV-specific outcomes. It was then attacked for using the word “conflict” rather than the word “violence,” just as the Crime Commission (1967b) had, as the political mood shifted toward punishment in the late 1970s.

**Arrests for misdemeanor domestic assault.** The trend toward demanding more punitive policing came to the Minnesota state legislature in 1979; the result was a new law allowing police to make misdemeanor arrests for domestic assaults they had not witnessed. This was not a law police wanted to use, so the Minneapolis Mayor and Council agreed to an NIJ experiment in which volunteer officers would randomly assign arrest, removal, and mediation (Sherman and Berk, 1984a). In a 6-month follow-up analysis, it was demonstrated that arrest worked best.

**Replications of the Minneapolis experiment.** Although this initial result had an enormous policy impact (Sherman and Cohn, 1989), the authors of the experiment had accompanied the findings with a strong recommendation for replications to be done as soon as possible. Five independent replications were completed, with varying results (Maxwell, Garner, and Fagan, 2002; Sherman, Schmidt, and Rogan, 1992). Even though the main effects of arrest versus nonarrest alternatives varied across the six cities, the most important finding was the consistent interaction between arrest and suspect unemployment. In experiments with three different demographic groups predominating in their respective samples, the effect of arrest was to deter employed offenders and increase repeat offending among unemployed offenders in three different ethnic groups: for African Americans in Milwaukee \( p = .006 \); Sherman, Schmidt, Rogan, Gartin, et al., 1992: 164), Hispanics in Miami \( p < .001 \)(Pate and Hamilton, 1992: 694), and Whites in Omaha (n.s. with smaller sample, but larger interaction effect than Milwaukee; see Sherman, Schmidt, and Rogan, 1992: 376–377); see also Berk, Campbell, Klap, and Western (1992).

The Milwaukee experiment also found that, even 24 years later, arrests of unemployed suspects for common assault caused significantly more domestic violence repeat offending than did an on-site warning (Sherman and Harris, 2013), with all arrested suspects more likely to be murdered than those who had been warned (Sherman and Harris, 2013). Most important, the all-cause mortality of victims was significantly higher if their abuser had been arrested versus warned (Sherman and Harris, 2015), with the effect particularly strong among African American victims (100% increase in death rates from 5% to 10% among victims, whose average age was 30 at the beginning of the 24-year period). This finding can and should be replicated in Charlotte, North Carolina, which had the only other predominantly African American sample among the six NIJ arrest experiments.

**Warrants for domestic assault.** Outcomes from another study recommended by the Crime Commission (1967b) produced very important results for the nearly half of cases in which offenders have left the scene before police arrive. Dunford (1990) produced strong
evidence that issuing an arrest warrant for an absent domestic violence suspect has a “sword of Damocles” effect in suppressing repeat injuries to the crime victims. In a randomized controlled trial in Omaha, Nebraska, Dunford enrolled cases in which suspects had already left the scene before police arrived. The police then randomly assigned either the standard police policy of advising the victim how to pay for an arrest warrant to be issued in court or went to the court to arrange for the warrant to be issued on police authority. The result was a 50% reduction over a 1-year follow-up of repeat arrest to the victims for whom police had obtained warrants against the offender, compared with those with no warrant, and a 40% reduction in injury to victims over 1 year. This experiment, however, also needs to be replicated.

“Second responder” follow-ups after domestic incidents. Several randomized controlled trials have tested follow-up visits by “second responder” police or social workers, or both, in a week or so after police have responded to a domestic violence call or made an arrest. A systematic review of these experiments included both repeat calls to police and interviews with the victims (Davis et al., 2008). The review found no difference between the victim-reported victimizations of those victims who had been visited and those who had not, but there were more calls to the police about new violence by those who had been visited than by those who had not been visited. Although the evidence is ambiguous, a second response cannot be considered beneficial.

Postarrest programs. Two recent postarrest experiments in England have found reductions in the harm severity of repeat violence based on randomly assigned police decisions. One is for first offenders; the other is for repeat offenders. For first offenders, an experiment in the port city of Southampton took place where police asked people with many years of experience in talking with domestic abuse victims and offenders to design a program to which they could refer first offenders. The program was for male offenders only, who were legally required by police to attend two weekend workshops, 4 weeks apart. The control group had no requirement except to avoid reoffending. The experiment was called CARA, enrolling almost 300 cases over 3 years. In a 1-year follow-up, offenders assigned to the workshop group were rearrested for crimes with a total CHI value that was 27% lower than the rearrests of offenders assigned to the control group (Strang et al., 2017).

For repeat offenders, a multi-agency partnership in Thames Valley focused on 180 couples who had generated two or more callouts by police within a 1-year period (Goosey, Sherman, and Neyroud, 2017). The couples were randomly assigned to receive either an “integrated case management” program (including services for both offenders and victims) or not. The control group had some contamination with victim services, but there was a major difference between the groups: Most treatment group offenders had repeated “checkup” visits by police compared with almost none in the control group. The outcome of the intent-to-treat analysis showed reduced harm but a higher volume of the repeat offending than in the control group.
Tracking
An early qualitative tracking of police compliance with mandatory arrest for domestic abuse in Arizona (Ferraro, 1989) found that police were generally ignoring the mandatory arrest law, with little consequence to them for doing so. In general, the United States has had almost no tracking research for three decades, whereas the English police are tracked (and criticized) regularly for not making enough arrests for domestic abuse. The recent development in the United States of the National Incident Based Reporting System (NIBRS) will also provide more incident-specific information of domestic violence incidents, thus allowing for better national tracking of domestic violence and its response compared with the Uniform Crime Reporting (UCR) system.

What Might a Similar Commission Today Recommend?
Mandatory arrest laws for misdemeanor domestic abuse were enacted in 28 states in the 1980s (Sherman and Cohn, 1989; Sherman, Schmidt, and Rogan, 1992), but it is not clear what difference this has made in criminal justice practices, let alone in outcomes for victims. A new commission today might recommend repealing those laws and exploring more promising alternatives, based on the last half-century of research.

Limiting the harm of arrest. The substantial evidence that mandatory arrest for misdemeanor domestic assault is harming African Americans is the highest priority issue in policing domestic abuse for a new presidential commission to consider. Strong evidence shows that arrest of unemployed African American offenders caused a 24-year increase in repeat offending, as well as an increase in the risk of the offenders being murdered (and not by their victims). It also shows that Black victims were twice as likely to suffer premature death if their offenders were arrested than if their offenders were warned. A new commission could do much to draw attention to these inequities and the need to correct them. It could also carry out the follow-up research on three randomized experiments to examine 30-year mortality rates for victims in relation to arrest of their abusers, to see whether the Milwaukee mortality findings are replicated elsewhere and continue over time.

Restorative justice. The findings from new research in the United Kingdom demonstrate that there is growing interest among domestic abuse victims to have an opportunity to participate, with extended family members, in a police-led restorative justice conference (Nettleton, 2017). These conferences have been generally effective in randomized trials for reducing repeat offending for other kinds of violence (Strang, Sherman, Mayo-Wilson, Woods, and Ariel 2013) and have been resisted for political reasons that may now be changing as U.K. police face major budget cuts.

Other postarrest actions. The two successful U.K. experiments testing workshops for first offenders, and repeat police checkups for repeat offenders, could be tested in the United States. So, too, could prosecution for the more serious cases that are nonetheless dropped if victims refuse to cooperate. Use of police body-worn video evidence has resulted
in successful prosecutions for serious cases in the United Kingdom and could be tested with controlled experiments in the United States.


This is what the Commission report said on p. 92:

A common kind of situation that illustrates the complexity, delicacy—and frustration—of much police work is the matrimonial dispute, which police experts estimate consumes as much time as any other single kind of situation. These family altercations often occur late at night, when the only agency available to people in trouble is the police. Because they occur late at night, they can disturb the peace of a whole neighborhood. And, of course, they can lead to crime; in fact, they are probably the single greatest cause of homicides. Yet the capacity of the police to deal effectively with such a highly personal matter as conjugal disharmony is, to say the least, limited. Arresting one party or both is unlikely to result in either a prosecution or a reconciliation. Removing one of the parties from the scene, an expedient the police often resort to, sometimes by using force, may create temporary peace, but it scarcely solves the problem. An order to see a family counselor in the morning is unenforceable and more likely to be ignored than obeyed. And mediating the difficulty of enraged husbands and wives ad hoc is an activity for which few policemen—or people in any other profession—are qualified by temperament or by training. Again no statistics are available, but there is a strong impression in police circles that intervention in these disputes causes more assaults on policemen than any other kind of encounter.

Since police action is so often so personal, it is inevitable that the public is of two minds about the police: Most men both welcome official protection and resent official interference. Upon the way the police perform their duties depends to a large extent which state of mind predominates, whether the police are thought of as protectors or oppressors, as friends or enemies. Yet policemen, who as a rule have been well trained to perform such procedures as searching a person for weapons, transporting a suspect to the stationhouse, taking fingerprints, writing arrest reports, and testifying in court, have received little guidance from legislatures, city administrations, or their own superiors, in handling these intricate, intimate human situations. The organization of police departments and the training of policemen are focused almost entirely on the apprehension and prosecution of criminals. What a policeman does, or should do, instead of making an arrest or in order to avoid making an arrest, or in a situation in which he may not make an arrest, is rarely discussed. The peacekeeping and service activities, which consume the majority of police time, receive too little consideration.
This is what the Commission report said on p. 104:

Handling minor disputes is an activity that is regarded as of small importance by most police administrators. Yet it occupies a great deal of the time of many policemen. To the disputants themselves, who are more often than not law-abiding citizens, the manner in which the police intervene in their affairs is a matter of great importance. Disputes, particularly domestic disputes, as discussed earlier, are a subject about which it would be difficult to formulate policy without first engaging in considerable research. The police should seek to accumulate information about families that cause repeated disturbances, to discover whether certain kinds of disturbances are more likely than others to lead to serious assaults or to homicides, to compile statistics on the typical effects of having one of the parties swear out a complaint against the other, to become familiar with the social-service agencies, if any, to which troubled families can be referred. For the police to mediate, arbitrate or suppress each dispute that they encounter as if it were unique—or as if all disputes were alike—contributes little, in the long run, either to law enforcement or to community service.

References


**Lawrence W. Sherman** is the director of the Jerry Lee Centre of Experimental Criminology at the University of Cambridge and a distinguished university professor of criminology and criminal justice at the University of Maryland—College Park.