HUMILITY IS A VIRTUE: ON THE PUBLICIZATION OF POLICY-RELEVANT RESEARCH

RICHARD LEMPERT

In their paper in this issue Sherman and Cohn (1989) respond to arguments I made in this journal about when research is ripe for publicity. This paper continues the conversation by pointing out that policy-relevant research may be publicized at three levels: (1) in professional journals, (2) directly to those practitioners whose practice decisions might be informed by the research results, and (3) through the mass media. I then argue that the reliability of the results, the ability to communicate main findings precisely, and the likely effects of publicity are keys to responsible publicization at all levels. In weighing these factors researchers must regulate themselves, but to guide them the law and society community should seek to develop professional norms. I comment on some of the considerations that should guide such norms, using the Minneapolis Domestic Violence Experiment and the publicity it received as an example.

I. INTRODUCTION

In 1984 I wrote that the results of the Minneapolis Domestic Violence Experiment,[1] as reported by Sherman and Berk (1984b), may have been "prematurely and unduly publicized, and that police departments that have changed their arrest practices[2] in response to this research may have adopted an innovation that does more harm than good" (Lempert, 1984: 509). These comments were not intended as criticisms of the study, for I thought and continue to think that the Sherman and Berk report is an excellent piece of work. Nor were they intended as criticisms of Sherman or Berk, for I did not know whether they had made any special efforts to publicize their results beyond publishing them in a scholarly journal, an enterprise that I not only regarded as appropriate but wished that this journal—which I was then editing—had been involved in. My comments were intended to raise, for the law and society community, the question of when research findings are firmly enough established to be "sold" to policy makers and to call

I would like to thank Shari Diamond for her editorial suggestions.

1 Henceforth "Minneapolis experiment."

2 I was thinking here of the establishment of mandatory or routinely arrest policies. The former, at least, Professor Sherman also opposed.

LAW & SOCIETY REVIEW, Volume 23, Number 1 (1989)

attention to the need to study empirically the process by which research results are disseminated. I am gratified that Professor Sherman and his colleague, Ellen Cohn, have seen fit to pursue these questions and so convert my statement to a dialogue, and I am grateful to Professor Diamond, the *Review*'s current editor, for her invitation to continue the conversation. I hope that other readers will join in.[3]

In their article in this issue Sherman and Cohn (1989) document the degree to which police departments learned of, understood, and were influenced by the Minneapolis experiment,[4] and they present the inside story of how the experiment's results came to be widely disseminated. They also argue that the publicity the Minneapolis experiment received was appropriate, particularly if the standards medicine applies in testing new drugs, which I saw as a possible model for law and social science, are any guide.

Sherman is to be commended for his foresight in attempting to trace the dissemination and influence of his work as well as for his candid account of how he sought to publicize his research results. We have so little empirical information about these processes that what Sherman and Cohn tell us makes a genuine contribution. Indeed, if Sherman's experience generalizes, the fear I expressed in my comment that policy makers will come prematurely to rely on single studies appears excessive if not downright silly. If work as well done, as timely, and as intuitively appealing as the Sherman and Berk study would, as Sherman and Cohn suggest, have received relatively little attention had there not been special efforts to draw attention to it (including such "Madison Avenue" techniques as direct "marketing" to likely implementors and releasing findings on a slow news day), sociolegal researchers are far more likely to find people ignoring work that deserves broad attention than to find an inappropriate reliance on work by policy makers or undue publicity in the larger community. Indeed, it appears that a major lesson to be drawn from Sherman's experi-

---

3 Some professional associations, like the American Psychological Association with its *American Psychologist* and the American Sociological Association with its *American Sociologist*, publish journals devoted largely to "issues of the discipline" of this sort. The Law and Society Association does not publish such a journal, and it is probably not economically feasible for it to do so. Thus discussions of disciplinary issues are largely confined to papers presented at the Association's annual meeting and, on rare occasions, articles in this *Review*. Perhaps the Association should consider expanding its newsletter to allow for brief statements addressing disciplinary topics.

4 Methodological problems mean that we cannot trust Sherman and Cohn's numbers precisely, but there is little reason to suspect gross inaccuracies in the basic picture they present. Methodological difficulties include the fact that respondents in different police departments occupied different positions and may not have been equally well situated to know whether and how their departments responded to the Minneapolis experiment, the possibility suggested in their Table 5 that earlier interviews affected later responses, and the possibility that some respondents who are scored as correctly identifying the findings of the experiment may have been merely guessing.

ence is that in all but the most exceptional instances it should be easy to limit the influence of research that suggests fruitful lines for policy reform but does not yet justify widespread reliance. If researchers conformed to professional norms about when research was unripe for popular dissemination, they would be unlikely to be thwarted by the media.

Sherman and Cohn are to be commended for pursuing and elucidating these important issues and for describing Sherman's publicity-seeking behavior in sufficient detail so that we may use him as an example. Sherman and Cohn argue for a position that justifies Sherman's publicity seeking; I shall argue, using the Minneapolis experiment as an example, for greater modesty and restraint. Indeed, one aspect of Sherman's publicity seeking provides what I see as an object lesson in the dangers of courting publicity.

II. TELEVISION FOOTAGE

Sherman, in an unpublished paper with Hamilton (1984: 3), wrote, "In order to keep a commitment to the Twin Cities public television station to supply findings for a documentary it produced on the experiment, the Police Foundation released preliminary findings in April of 1983. The final report was not published until April of 1984." I was troubled when I read in Sherman and Cohn that Sherman had persuaded a Minneapolis public television station to produce a documentary on his research with the idea that this would enable him to provide action film footage to other television shows when the experimental results were later released.[5] But it is even more troubling when the price of securing such coverage is a promise that research results will be released by a specified date, one that may be long in advance—in the instant case it was a year—of the date when the data would otherwise have been sufficiently well analyzed to be released to the scientific community. Moreover, Sherman not only released data that had not been analyzed up to the professional standards that he and Berk later insisted on, but he also sought to publicize the data nationwide. In publicizing these preliminary findings Sherman was careful to mention reasons why they had to be treated cautiously by policy makers, but it was predictable that many of his cautions would be lost as the study's results were disseminated through the media.[6]

When a desire for publicity accelerates the schedule on which

[5] This situation would be different if a researcher had persuaded a television station to produce a documentary on field research in order to show how social science research is conducted. Here the researcher's incentives are to show how proper research is conducted and not to ensure that vivid footage—suitable for later "highlight films"—will emerge. Of course a television station would be unlikely to find social science research worthy of a documentary unless exciting scenes could be expected. Thus, a researcher who invites such coverage might be pressured to distort the design or conduct of the research to make for better "photo opportunities."

[6] An exception was Ellen Goodman's column (1983), which was remarka-

social scientific data are released, as apparently occurred because of Sherman's commitment to Twin Cities Television, the standing of the relevant social science may be harmed, and bad advice may be given to policy makers.[7] Sherman's situation, I think, serves to warn of dangers that can arise when decisions about the conduct of research and data analysis are tailored to considerations of publicity.[8] It also suggests a need for professional norms that preclude such commitments.

III. PUBLICITY AND AUDIENCE

Sherman and Cohen defend Sherman's actions with respect to the Minneapolis experiment and appear willing to generalize from his behavior to a standard for the profession. I continue to think that the results of the Minneapolis experiment were unduly publicized, even though I admit, as I have elsewhere (Lempert, 1987), that if I were a police chief with a departmental policy against arresting in misdemeanor spouse assault cases, I would modify that policy in light of Sherman and Berk's findings. This admission is not inconsistent with my position on what the norms of social science should require.

In thinking about the degree of publicity that is appropriate for policy-relevant social science, it is helpful to distinguish among three levels on the basis of audience. The first is publication to the specialized scientific community. The second is dissemination to the professionals who are most likely to apply the research results. The third is dissemination to the general public. The levels differ in the sophistication of the audience that receives the message, the

---

bly sensitive to the reasons why the data Sherman had released were interesting but might not justify widespread policy changes.

[7] In the instant case these dangers were apparently not realized. As Sherman interprets them, the policy implications of the fully analyzed data are not very different from the implications of the preliminary analysis, but the possibility that the early conclusions would have been called into question by more refined analyses was there.

I am not saying that it is inappropriate to publish the results of ongoing research or less-than-fully-analyzed data sets. This is often done, both because of academic pressures to publish and to establish priority for novel ideas. But decisions regarding ripeness for publication should not be corrupted by commitments that have been made to assure publicity for one's research. There is also a difference between publishing results for professional audiences and publicizing them to a lay populace. Professional audiences are more likely to appreciate the limitations of the analyses in light of what is to follow, and if later research requires the revision of earlier conclusions it will be easier to communicate this fact to narrow professional than to widespread lay audiences.

[8] I am not suggesting that Sherman entered into the agreement with Twin Cities Television intending to publish his data before they could be fully analyzed. When he made his commitment to Twin Cities, Sherman may have anticipated that the final results of the data analysis would be available long before the television station would want them. But research often takes longer than anticipated. A danger of courting publicity is that it may lead one to make commitments that ignore or minimize the likelihood of such contingencies.

information that is conveyed, the likely policy impact, and the costs and benefits of publicization.

In my editorial comment (1984) on Sherman and Berk, I did not say nor did I imply that their study should not have been published in the *American Sociological Review*. Indeed, I said that as the editor of the *Law & Society Review*, I wished I had been able to publish their article in this journal. Professional journals allow researchers to communicate their findings in their own words to knowledgeable audiences. Peer review provides some guarantee that policy-relevant conclusions are supported by properly analyzed data. Technical problems and their resolution may be explained to readers, at least some of whom will be sufficiently knowledgeable to evaluate the author's methodology. Weaknesses in theory or analysis may well lead to published comments that will alert essentially the same audience that read the original study to controversial aspects of the research results. Moreover, even if there is a risk that a study's findings will be publicized beyond the research community—a risk that appears, from Sherman and Cohn's discussion, to be small—suppressing publication in professional journals for reasons other than peer judgments of quality poses a serious threat to core norms of science. Thus there is no doubt in my mind that the publication of Sherman and Berk's study in the *American Sociological Review*—even when it stood almost alone as a study of the police control of spouse abuse—was entirely appropriate. Indeed, publication of their study would have been appropriate had its results been either more equivocal or more likely to affect policy than they were. Science cannot grow unless studies like this are published.[9]

At the second level is dissemination of research results to groups of professionals or semi-professionals who administer policies to which the findings relate. The decision by the Police Foundation to mail a simplified version of the study report to 3,000 policy makers (most of whom, I assume, were police officials) exemplifies dissemination of this sort. The appropriateness of dissemination at this level will vary with the quality, reliability, and policy relevance of the research disseminated. Clearly research is much more likely to have a policy impact if its findings are distributed directly to a relevant set of policy actors than if they are reported only in a professional journal. At the same time the simplification needed to reach the wider audience may ignore or gloss over factors that have policy relevance. In particular, readers of

[9] I am not saying that it is never appropriate to suppress the publication of research at this level. This is a question I do not address except to note that there are values other than scientific ones, and that even for scientists it is hard to make an ethical case that, when values conflict, scientific ones must always predominate. However, for me at least, the Sherman and Berk study is far removed from those extreme situations in which the ethical case for the suppression of scientific publication has any appeal.

simplified research reports will be less able than readers of the underlying research to put their own interpretations on the reported data, and they are less likely than scientific audiences to become aware of criticisms of the results furnished them.

On the other hand, when researchers address policy makers directly, they can largely control the message conveyed. If the researchers are as knowledgeable and as professionally honest as Sherman and Berk, they can be sure that their readers are alerted to limitations in their data and to the possibility that their results will not generalize. Moreover, target audiences will contain people of differing levels of sophistication. Some will go back to read the original research reported in the summary, especially if their agencies are considering new policies based on it. Other individuals or agencies may direct questions to the researchers or other experts before they rely on results summarized for them.

Had I conducted the Sherman and Berk research, I do not believe that I would have mailed reports of the study to 3,000 policy makers with the aim of influencing their action. I would have preferred to see the results replicated before taking this step. Nevertheless, I make no strong claim that the Police Foundation's actions in this respect were inappropriate or that the standards of the scientific community should preclude this. A good case can be made that information from methodologically sound studies should be fed quickly into relevant policy communities to enter into the ongoing discourse about how to deal with social problems. Indeed, Sherman and Cohn argue that without mass media publicity it is difficult to move action agencies to reforms that reflect the latest social science findings. Thus the problem that most concerns me, which is that research results may lead to policy changes that are more harmful than helpful, is unlikely to occur. Certainly I would not have written my editorial comments had Sherman done nothing more than send summaries of his study with Berk to police and other crime control policy makers.

This brings us to the third level of dissemination, which is publicity through the mass media: the press, television, and the like. Here the researcher can be confident that cautions carefully attached to a scholarly report or practitioner summary will, in many instances, not be communicated alongside the basic research findings. The researcher can also predict that groups that are fighting for "reform" for reasons that have nothing to do with the empirical situation will seek to make political hay of any results that prove congenial. Finally, as Sherman and Cohn argue, policy change becomes more likely not because the relevant policy maker has evaluated the research and thinks change is wise but because political and editorial forces build up that make change irresistible.

Thus even if researchers should not be required to take account of the possibility that general publicity may attend the dissemination of results at the first or second levels I have described, researchers who actively promote the dissemination of their findings in the mass media—who, for example, release their results on slow news days or who provide television news programs with film clips from their research—bear, in my opinion, a heavy burden. The burden is one of reliability; before promoting research through the mass media researchers should be confident that the results they report are sufficiently reliable and generalizable, even without appropriate limitations and cautions, to justify the impact that they may have should the publicity efforts succeed in full measure. Even if researchers have faith in the conclusions they would draw from their work, they must consider the likelihood that precision will be lost as findings are disseminated through the mass media and that the implications they would derive from their work will not control its uses in the political process. For example, in reporting his results to the media Sherman made it clear that he did not believe that his findings justified a policy of mandatory arrest in all cases of misdemeanor spouse abuse. Yet his research has apparently been a factor leading some jurisdictions toward a mandatory arrest policy.

A good case can be made that the burden to be met before deciding that policy-relevant research deserves mass media dissemination is so great that a researcher—who may think better of his work than others do or may find mass media publicity personally gratifying—should never decide whether his own work deserves such attention. Rather we should have a form of peer review or a science court that acts as a gatekeeper to the mass media. However, we currently lack such an institution,[10] and there is little prospect that the social sciences will become disciplined in this way in the near future. Thus the most that we can do is share ideas on the matter, work to develop a normative consensus that can guide researcher decisions, and help politicians and the media to become more educated consumers of social science.

---

[10] We do have some institutions now that perform this function at a more general level, most notably the National Research Council (NRC), which is the research arm of the National Academy of Sciences. However, the NRC rarely conducts original social science research, nor does it clear single studies for dissemination to the media. Rather it reviews work in an area and attempts to sort out what social science may reliably tell us from what it cannot tell us and from the downright misleading. In the course of doing this, individual studies may be singled out for praise or criticism. NRC panels contain a mix of views, and authors of celebrated research that will figure heavily in the panel's deliberations are often excluded for just that reason, although they may be given a chance to appear before a panel to discuss their findings.

Still more generally, scientific communities filter research for the media. Well-supported findings become common knowledge in the disciplines, and these disciplinary perspectives are communicated in various ways to the media, such as through commentaries on current events. For example, there is now widespread popular knowledge that fluorocarbons deplete the ozone layer; that AIDS cannot be spread through casual contact, and, in a law-related area, that drunk drivers are responsible for a large proportion of fatal auto accidents.

IV. A STANDARD OF PUBLICITY

Clearly the fundamental consideration concerns the reliability and generalizability of the results the researcher might disseminate. It is here that Sherman and Cohn and I differ with respect to both the question of whether the Minneapolis experiment was of a quality to justify the publicity Sherman sought and the issue of what general standards in the area should be.

In their article Sherman and Cohn are candid and fair in disclosing shortcomings that might lead one to hesitate before generalizing from the Minneapolis experiment to policy in other areas. I might simply refer the reader to their discussion and say, "Case made." However, there are two points that deserve further attention.

The first, which has been generally overlooked, is that there is a fundamental ambiguity in the Sherman and Berk results. They measure the relative effects of arrest, counseling, and separation in two ways: through subsequent complaints to the police and through interviews with those women who were willing to meet with project interviewers for up to six months. By the first measure arrest is a reliably better treatment than separation but is not reliably better than counseling at conventional levels of statistical significance. By the latter measure arrest is reliably more effective than counseling but not more effective than separation.

Now my hunch is that these differences simply reflect a lack of power in the statistical tests; that is, I expect that arrest would have been reliably superior to each of the other treatments by both measures had more cases been included in the sample. Yet nothing but a theoretical leap—my own faith in arrest based on little evidence apart from this study—justifies this conclusion. Recognizing this, the matter is troublesome in a study that has had great policy impact. In particular, as I pointed out in my earlier commentary, this pattern of results might also be found if arrest did not deter men from beating their spouses but instead deterred women who were beaten from calling the police.[11]

Moreover, there are potentially serious problems with each of the study's measures of repeat spouse abuse. The measure based on interview responses suffers from substantial nonresponse in a setting that is fraught with the possibility of response bias. The measure based on calls to the police is rendered somewhat ambiguous by the fact that many calls come not from the abused spouse

---

[11] Sherman and Cohn acknowledge the possibility of intimidation but suggest it is unlikely because the response rates of interviewees did not vary by treatment and that, given the ratio of victim to nonvictim calls to the police in the experimental period, intimidation would have to have entirely eliminated victim calls during the follow-up period to yield the data found. Sherman and Cohn's latter claim is misguided. Only some victim calls would have had to have been eliminated to yield results that would have invalidated the claim that there was a statistically reliable difference between treatments.

but from others.[12] Thus it may be that the police-based measure to some extent reflects a reduction in behavior that is sufficiently noisy to bother neighbors but that is not regarded as abusive by the spouse concerned.[13] These possibilities do not undercut the scientific merits of the study, but they do suggest good reason to refrain from a media campaign.

My second point concerning the Sherman and Berk study relates to what Sherman and Cohn call the "micro-mediation" issue. In the April 1984 Police Foundation report on the Minneapolis experiment (Sherman and Berk, 1984a) we learn that, according to interviews with the 49 percent of the victims who could be followed for six months, when the police not only arrested the abuser but also took the time to listen to the victimized spouse, the recidivism rate was about 9 percent.[14] When the police arrested the abuser but did not attend to the victim, the recidivism rate was 26 percent. It appears from these data that the statistical significance of the difference between the arrest and counseling treatments as reported by victim interviewees is due to the subgroup of cases in which the victims not only saw their abusers arrested but also felt that the police were interested in them. Had these victims experienced subsequent abuse as frequently as those who only saw their abusers arrested, the victim interviews would have yielded no reliable evidence that arrest was superior to the other treatments. But this kind of contextual consideration was, predictably, almost entirely lost in the media dissemination of the experiment's results.

If attention to the victim is a crucial mediating variable,[15] reforms that increase arrests in spouse abuse cases could do more

---

12 Sherman tells us that abused women made only 45% of the calls to the police during the experimental period. Data on who called the police are not available for the follow-up period.

13 Evidence from the interview sample of the association between women defined as victims by self reports and those defined as victims by police reports would address this possibility, but it is not provided.

14 The figures are only approximate. The Police Foundation report presents not raw data but adjusted estimates from a multivariate model that included the effects of the prior number of arrests for crimes against persons.

15 Sherman and Cohn state that, according to a private communication, Berk came to believe that the difference between the arrest plus attention and the arrest plus no attention conditions that I have described was not sufficiently reliable to justify a paper exploring the interaction. I have the greatest respect for Berk as a methodologist, but the difference given in the Police Foundation report is so striking that a more detailed explanation of why this is unimportant than that which Sherman and Cohn provide appears warranted. Even if attention to the victim does not matter, the points made in the text still hold, because the *American Sociological Review* article (Sherman and Berk, 1984b) and the Police Foundation report (Sherman and Berk, 1984a) both appeared in April 1984, shortly before Sherman's efforts to engineer the second wave of publicity commenced. Thus the significance of the fact that an interaction effect with crucial policy significance was underemphasized or missed in much of the publicity the study received is not vitiated by the fact that later data analysis showed the effect not to be important, assuming this is the case.

harm than good. Under unreformed systems a victim's preferences—which can only be determined by attending to the victim—are often an important factor in the police decision to arrest. A policy of routinely arresting might make the police less attentive to the victim, since her preferences will not determine their arrest decision. The important point is not, however, that increasing the rate of spouse abuse arrests will not have the desired deterrent effect. Rather it is that social science research findings are often contextually contingent in ways that bear importantly on policy implications. To attempt to influence top policy makers through the political pressures the media can generate carries a substantial risk that information about important contingencies will be lost. To attempt to build a scientific consensus about what matters and to present this consensus to policy makers directly or through existing agency channels is a slower process than the dissemination of results through the mass media, with fewer rewards to researchers who find the limelight gratifying, but it is more likely to lead to sound policy that wisely uses what social science can tell us.

V. MEDICAL TESTING AND REPLICATION

In my comments on the Sherman and Berk study, I suggested that their work should have been replicated before it was published in the media, and I suggested that social science might take a cue from medicine with respect to the kind of testing that should be required before solutions to social ills are claimed. Sherman and Cohn respond by arguing that in medicine studies ordinarily need not be replicated before the results reach the media and influence treatments. Sherman and Cohn write as if the truth of their argument would refute me. However, they misunderstand my argument from the analogy to medicine.

I did not say that the widespread publicity accorded the Sherman and Berk study should have followed rather than preceded replication because medicine routinely replicates its treatment research. I said rather that in medicine new drugs, even very promising ones, are painstakingly tested before they are approved for distribution. What constitutes adequately painstaking testing depends on a treatment's possible side effects, the condition it is intended to alleviate, the theoretical basis for the treatment, and the quality of the research that supports the treatment. Moreover, standards for publicizing the results of medical studies in the mass media may, perhaps, be properly more relaxed than the standards for publicizing the policy-relevant results of comparably reliable social science investigations. Whatever the popular impact of a medical success story, the publicized treatment will often be administered by professionals who can appreciate the more complex and technical story that underlies the publicized version. Indeed, even if a treatment does not require professional administration—as with taking aspirins to prevent heart attacks—people nevertheless may consult with physicians. Certainly legislatures are not going to mandate that every adult take an aspirin a day because a study finds this beneficial.[16]

Sherman and Berk's study should have been replicated before its results were disseminated to the mass media not because replication is *de rigueur* in medicine, but because legal and social conditions in Minneapolis are sufficiently different from the conditions in other cities so that what works there (assuming arrest does work) may not work as well or in the same way in other cities. A researcher who helps foster nationwide pressure for a reform should have good reason to believe that the results supporting that reform are generally applicable. The problem is not that the Minneapolis experiment was a single study; the problem is that it was a *Minneapolis* experiment. Had Sherman and Berk designed a study that collected data from ten cities simultaneously and had the results been consistent across locations, I would not have called for replication—single study or no—and I might not have questioned the publicity.

One reason I caution generally against great reliance on single studies is that most studies are in some ways narrowly confined and leave open a number of credible threats to external validity.[17] Another reason is that our experience with single studies that have received extensive publicity for their general bearing on policy problems is a humbling one. In their introduction Sherman and Cohn cite a number of studies that received substantial public-

16 When professional mediation is unlikely, I think medicine should demand a higher standard of reliability for studies publicized in the media than it demands when professional mediation is likely. I do not know whether such an enhanced standard is required.

17 Sherman and Cohn miss this point when they state that single studies often lead to medical innovation. Medicine has long experience with drug tests conducted on nonrandom samples of individuals that generalize to the larger population. Even if there is some debate on this point with respect to certain studies within medicine, for most purposes different individuals may be assumed to be more similar in their reactions to drugs than different cities may be assumed to be in their reaction to social innovations like a routine arrest policy for misdemeanor domestic violence. The difference lies at the level of theory. In the case of many drugs, animal testing and biochemical theory may mean that there is little reason to expect different populations to respond very differently even if some people, such as an allergic subgroup, have responses that differ from modal or mean reactions. When there is reason to expect different reactions, as with young children or pregnant women, further testing must be done on samples from the particular population before the drug may be relied on for them. In the case of the Minneapolis experiment, there are plausible theoretical reasons to expect that the implications of an arrest policy will be affected by characteristics that can vary from city to city, such as post-arrest treatment, unemployment, ethnicity, climate, and police professionalization. Thus replication is needed before the Minneapolis arrest policy can be confidently recommended to other cities, just as replication of a test done on males may be needed before a drug can be regarded as safe for pregnant women.

ity for findings that now appear suspect and I could name others, but they do not draw from these studies the lessons I would for the issue we are discussing.

VI. THE POINT OF THE ANALOGY TO MEDICINE

I intended the analogy to medical research not to justify my call for replication but to support a far more difficult point. The strongest part of Sherman and Cohn's position and the weakest part of mine is that the status quo is itself a treatment. In jurisdictions that lack an arrest policy for misdemeanor domestic violence, women are being abused by their husbands or lovers every day, and it is possible that such abuse will escalate into more serious or even fatal violence. To refrain until all the results are in from publicizing, in the most effective way possible (that is, through the mass media), research that suggests that arrest deters domestic violence may mean that women are abused or even murdered because their efforts to invoke police protection do not result in arrest. Given these possibilities, how can we not publicize research that suggests that changing the status quo is desirable?

It is for those who think that the answer is obvious—that we clearly should not refrain from publicizing results in these circumstances—that I refer to medical practice. In clinical trials doctors may withhold a new drug with beneficent or even lifesaving potential from a control group in order to better ascertain the true effects of the drug while also identifying unwanted side effects. Such experiments may continue past the point where it appears the drug is likely to be safe and effective in order to establish safety and effectiveness with greater certainty. Sherman and Cohn respond by citing the well-known aspirin and AZT studies, each of which was terminated before its planned conclusion so that the news of treatment effectiveness could be widely disseminated and people could benefit from it. However, these studies support rather than disprove my point. The aspirin study was a large-scale project that was terminated only when the results appeared reliable and replitable. Indeed, the aspirin study continued long enough with a large enough sample to pinpoint a rare but important side effect—an enhanced probability of strokes—and to document that the likely benefits from the prophylactic use of aspirin far outweighed the possible costs. The AZT study also continued past the point where it seemed that the drug was more effective than the placebo, and terminating the study as soon as the drug's effectiveness was clear was justified because continued placebo treatment meant certain death.

Now it might be argued that medicine should not be our guide in this area, or at least that the popular image of drug testing procedures that I present should not be a guide.[18] Indeed, there is

---

[18] I have presented an image of medical procedure based on casual read-

considerable controversy about whether the procedures I outline have unduly delayed the introduction of new drugs and valued therapies at great cost to the sick and injured. I believe, however, that there would be general agreement that, when deciding that a new drug has been adequately tested, the less serious the condition treated and the more serious the potential side effects the more certain we should be that a treatment works and that the actual side effects are rare or not harmful. While the Minneapolis experiment suggests that arrest does more to prevent subsequent domestic violence than the other treatments evaluated, it is not clear that this apparent finding is ripe for translation into policy when weighed in the "risk-benefit" balance.[19]

Without meaning to denigrate the harm prevented, we are talking about arrests for misdemeanor domestic violence rather than for felony assaults. Moreover, it appears the recidivist domestic violence that arrest apparently prevents is largely of the misdemeanor rather than the felony variety. Indeed, the recidivist violence measured by Sherman and Berk (1984b) did not necessarily involve later physical harm to the victim because subsequent property damage and threatened violence were counted as violence in the victim interviews, and behavior sufficiently disturbing to induce neighbors or others to call the police was counted as recidivism in the police records measure, even if it did not go beyond shouting. On the other hand, a number of costly side effects plausibly associated with arrest are not ruled out by the experiment's results. The reported data are consistent with the possibilities that the arrest of a spouse or lover deters women from seeking police aid in subsequent assaults and that the apparent deterrent effect of

---

ing and occasional news stories rather than systematic research. It is possible that this image is not how medicine operates at all, although I do not read Sherman and Cohn as disputing my portrait of medicine in this particular. It does not matter if I am accurate, for I am not suggesting that we should consider emulating medicine because it is medicine. Rather we should emulate medicine only if its procedures are sensible. Because medicine has flourished using procedures like those I describe at least with some treatments, there is at least some reason to doubt whether immediately publicizing social experiments suggesting that some treatment is better than the status quo will be good for social science and good for humanity. Perhaps, like medicine's norm of aspiration, we should have as a fundamental precept, "Do no harm."

19 Note that a good case can be made that spouse abusers should be arrested because justice demands it. One who holds this position may see little reason to withhold the results of a study like the Minneapolis experiment pending further testing, for disseminating the results can only help bring about a more just state of affairs. But such a person should find the results of the Minneapolis experiment irrelevant because he or she believes arrest is the appropriate treatment for value reasons that have nothing to do with the effects of arresting spouse abusers. However, my discussion proceeds, as does, I believe, Sherman and Cohn's, by bracketing the justice issue and focusing entirely on the empirical question. The "good" we are both concerned with realizing is minimizing domestic violence.

arrest exists because arrest breaks up relationships regardless of whether victims want them broken up.[20]

Other possible costs of arrest were not measured in the study. These include the possibilities that those arrested men who are not deterred from further violence are more violent than they would have been had they not experienced arrest; that arrestees may lose their jobs or opportunities for future positions thus harming both them and their victims; that arrestees may be victimized by jail violence after their arrest; and that arresting spouse abusers requires the allocation of valuable police time and jail space that might more effectively reduce social violence if deployed in other ways. Clearly even if arrest works as Sherman and Berk's study suggests, further testing is needed to evaluate the benefits that are traceable to arrest's deterrent effect and the costs that may attend it.

VII. CONCLUSION

Even if the Minneapolis experiment's results are accepted and even if there were no concern that external validity considerations might mean that arrest was not the most effective treatment in other jurisdictions, all jurisdictions would still have reasons to refrain from changing their arrest policies pending further investigation. Indeed, in medicine the degree of benefit coupled with the list of unexcluded possible side effects might demand that treatment implementation await further testing. Nevertheless, I have not suggested that the Sherman and Berk study should not have been published, nor have I said that its results should not have been distributed to police professionals, although I am doubtful on this score; I have not even said that it would be inappropriate for police departments to modify their arrest policies in light of the study's results.[21]

I *have* said that Sherman's special efforts to publicize the Minneapolis experiment's results are not in keeping with what I think good social science practice entails. This is based on what I have read of the experiment and of the publicity it has received.[22] I take this view not only because cautions essential to the proper in-

---

[20] Arrested men take longer to return to relationships with their spouses or lovers than those given the separation treatment. According to the initial victim interviews, 32% of the arrested men had not returned to the relationship within a week, and they may not have returned at all, while only 10% of the separated men fall into this category. Men who were counseled did not leave the relationship, but the violence that led to police intervention seldom if ever resumed when the police departed (Sherman and Berk, 1984b: 268).

[21] Indeed, I have stated that if I were a police chief I would change a "do not arrest" policy because of the study's results, although I would not mandate arrest (Lempert, 1987).

[22] I am grateful to Larry Sherman for conveying to me, through Shari Diamond, copies of news stories and editorials commenting on the Minneapolis experiment.

terpretation of data were predictably lost when the mass media reported the results of the Minneapolis experiment for general consumption, but also because the mass media publicity—also predictably—may have created a force for change that legislators or administrators, whether sensitive to questions the Minneapolis experiment left unanswered or unaware of them, found impossible to resist. It is the creation of such a partially informed force for change[23] that I see as the principal vice in going to the media with a study of doubtful external validity that leaves vital questions unanswered. Sherman and Cohn (1989: 141), on the other hand, seem to think that the creation of such a wind for change is perhaps the major virtue of publicizing a study's results in the mass media. They write:

> Mass media often fail to convey the full complexity of a study's findings, which readers need to interpret the results properly. But the media are the only form of publicity that can be sure to reach top policy makers, since it is the media that shapes the editorial and political pressures to which they must respond. To advocate publicity solely through professional channels may be to advocate burying research results so that they can have little useful effect on either current practice or the conduct of replications.[24]

I would have to have considerably more faith in the external validity of a study and in the overall beneficence of the treatment it favors than I do in the case of the Minneapolis experiment to think that the kind of popular publicity that Sherman sought was justified. Maybe I would have such faith if the Minneapolis experiment were my study! This is precisely why the law and social science community should continue the debate that Sherman and Cohn and I have begun and why the community should attempt to form a normative consensus on—or even an institutional way of dealing with—the publicity issue.[25]

---

[23] The force could just as well be one favoring the status quo, depending on the research results.

[24] Except in one particular, which I address in the following footnote, I fail to see why confining research reports to professional journals means that the research will have little effect on the conduct of replications.

[25] There is one point that Sherman and Cohn make that I have not dealt with in this response: that funding agencies might not be willing to support replications absent the pressures and satisfactions that attend widespread publicity. If this is the case there is a scientific justification for seeking to publicize through the mass media research that is otherwise not ripe for popular dissemination: The replication research needed to refine a preliminary understanding of a problem will not be supported without such publicity. I cannot dispute Sherman and Cohn's suggestion that, but for the publicity accorded the Sherman and Berk study, the National Institute of Justice would not have supported replications, except to say that my contacts with that agency suggest that their leaders and administrators are aware of the importance to both science and policy of replicating research like the Minneapolis experiment.

If Sherman and Cohn are right in saying that studies must be publicized to shake loose funds for replication, the social science community should collectively seek to educate agencies—particularly "mission" agencies—about the

In criticizing Sherman and Cohn's argument I have suggested that the widespread popular publicity that the Minneapolis experiment received was premature and that some of Sherman's efforts to secure such publicity were inappropriate. Obviously Sherman disagrees. Thus the reader is left to evaluate our competing arguments and to weigh in with his or her own contributions.

At the same time I hope that my criticisms have not obscured the important contribution that Sherman, writing with Cohn, has made. First, he provides us with rare, if imperfect, data on how knowledge of a celebrated social science study diffuses and comes to influence important policy decisions. Second, his even rarer candor about his efforts to publicize his research have provided us with the concrete example that makes this discussion meaningful.

I have also questioned the wisdom of the widespread policy changes that were motivated in part by the findings of the Minneapolis experiment, a research effort that I applaud as an excellent example of policy-relevant law and social science. The fact—if it is a fact—that reliance on the Minneapolis experiment has been premature should not be taken to suggest that law and social science research has nothing to offer the policy maker. Quite to the contrary; replication of the Minneapolis experiment is necessary because we potentially have much to offer. Seventeen years ago Donald Black (1972: 1100) concluded an article in the *Yale Law Journal* by writing, "At present, applied sociology of law has little to apply. What more serious claim could be brought against it?". That Sherman and Cohn and I should have this discussion is an indication of how far we have come.

REFERENCES

BLACK, Donald (1972) "The Boundaries of Legal Sociology," 81 *Yale Law Journal* 1086.

GOODMAN, Ellen (1983) "Using 'Muscle' Against Wife Beaters," *Washington Post* (April 19).

LEMPERT, Richard O. (1984) "From the Editor," 18 *Law & Society Review* 505.

LEMPERT, Richard O. (1987) in R. O. Lempert and C. A. Visher (eds.), "Randomized Field Experiments in Criminal Justice Agencies: Workshop Proceedings," Presented to the National Institute of Justice, National Research Council, (August) pp. 174–187.

SHERMAN, Lawrence W., and Richard A. BERK (1984a) "The Minneapolis Domestic Violence Experiment," April, *Police Foundation Reports*.

——— (1984b) "The Specific Deterrent Effects of Arrest for Domestic Assault," 49 *American Sociological Review* 261.

---

situations in which replications are essential to determine the parameters of the policy relevance of promising initial studies. This should be done not just to enhance the value of social science findings but also to prevent the widespread adoption of policies that turn out not to work or to backfire in particular settings. In the long run failures of this sort are likely to do more to reduce the willingness of agencies to fund social science research than momentary publicity will bring.

SHERMAN, Lawrence W., and Ellen G. COHN (1989) "The Impact of Research on Legal Policy: The Minneapolis Domestic Violence Experiment," 23 *Law & Society Review* 117.

SHERMAN, Lawrence W., with Earl HAMILTON (1984) "The Impact of the Minneapolis Domestic Violence Experiment: Wave I Findings;" Presented to the Police Foundation (April).