Perhaps overgeneralizing from impulses that occasionally move me, I have often thought that much of the research in law and social science is stimulated by the "urge to be relevant." Researchers are drawn to problems in our discipline not only by the scientific challenge these problems pose but also, and sometimes primarily, by a desire to speak to important social and political issues in ways that will directly influence official decision-making. A corollary of this is that signs of influence are often exaggerated both by those whose work is noticed and by the profession. Thus, much has been made of footnote 11, the "social science footnote," in *Brown v. Board of Education* even though there is no reason to believe that the evidence swayed any votes, and the social science is, at least in retrospect, crude and unconvincing. In a personal analogue, I recall one social scientist who pointed with obvious pride to a Supreme Court citation to his work when it was not only clear that this work had no effect on the decision reached, but it was also obvious that the Court did not understand what he had done.

These examples do not mean that social science has no influence on legal decision-making. One need only consider the movement away from the rehabilitative ideal toward determinate sentencing to realize that, in conjunction with other factors, this influence can be profound. But usually, to be influential social science learning must, as Harry Kalven put it, "become *popular* learning" (1968: 68), which ordinarily requires a body of learning that consistently points in one direction. Most articles that end, "It follows from these results that legal decision-makers should . . ." are consigned, and usually properly so, to practical oblivion. But by the same token, well-established findings are often ignored.

Occasionally, however, there is an article or book which draws such media attention that it almost overnight becomes popular learning and is seized upon as a prescription for social action. Such responses to single studies are almost always as inappropriate as the law's failure to consider what is established learning. Both suggest a popular misunderstanding of how social science proceeds and a failure on the part of social scientists to communicate a sense of the strengths and limitations of what we are about. The danger

LAW & SOCIETY REVIEW, Volume 18, Number 4 (1984)

that one study will be unduly influential is, I think, particularly acute in law and social science because much of our research relates directly to issues that are constantly being considered by legal actors and because we consciously seek to be relevant.

Recently, there has been an example that to my mind nicely illustrates the problems of which I speak and so raises a number of fundamental issues for our discipline. It is a study of police responses to domestic assaults by Lawrence Sherman and Richard Berk (1984). Basically, the study finds that the arrest of an abusing male appears more likely to prevent future assaults than either counseling or separation without arrest. Without making any systematic attempt to canvass this study's influence, I have seen its results mentioned in local, regional, and national newspapers; I have tuned in an ABC Nightline feature in which it figured prominently; I have heard it was a topic on the Today Show; I have read that both Minneapolis, Minnesota, and Newark, New Jersey, have dramatically increased their arrests in domestic violence cases in response to this research; and I have heard that similar changes have occurred in other jurisdictions as well.

The Sherman and Berk article was not submitted to the *Law & Society Review*, but I wish it had been, for it is a fine piece of work and I would have been delighted to publish it. The authors attempted and in large measure succeeded in implementing a true field experiment in the treatment of "spouse abuse."[1] Minneapolis police officers were persuaded by the experimenters to deal with misdemeanor domestic assaults according to the randomly ordered instructions on a report form. A six-month follow up, which included both victim interviews and a canvassing of police reports, revealed that by one measure (police reports) arrest was significantly more likely to eliminate subsequent violence than separation without arrest, and by another measure (victim reports from the 49 percent of all victims who could be traced for six months) arrest was significantly more likely than counseling to reduce repeat offending.

While the officers did not always follow the experimental instructions, they did so often enough that the strength of the experiment was not lost. More importantly, Sherman and Berk are properly sensitive to the breakdowns in the experiment, to

[1] Only 35% of the alleged assailants in the sample were husbands of those they assaulted. Forty-five percent were unmarried lovers of their victims, 3% were divorced or separated husbands, 15% were male relatives, roommates, or others, and 2% were wives or girlfriends.

problems with sample attrition and the like. Through the use of modern statistics and sound common sense they take account of many of the methodological issues that arise. They are similarly aware of a number of factors they could not control, and although they conclude by suggesting that in domestic assault cases a presumption of arrest is desirable, they also note a number of reasons why one might be cautious about accepting or generalizing from their results. No doubt, the quality of this study contributed to its ready acceptance. Indeed, it is because of its scientific merits that this study so nicely illustrates the dangers of generalizing from a single investigation.

These comments are not a "Response to Sherman and Berk," and I don't intend to discuss in detail every objection that may be made to generalizing from this research. It is, however, instructive to examine some salient objections. The most obvious is one the authors warn us about: "Minneapolis is hardly representative of all urban areas" (1984: 269). The racial composition of Minneapolis, the quality of its police force, and numerous other factors distinguish it from other cities. Perhaps most important is the fact that only three of the 136 arrested offenders in Minneapolis were formally punished by fines or subsequent incarceration. In another city with different prosecutorial policies a marked increase in domestic violence arrests might lead to a marked increase in prosecutions, with costs to offenders and perhaps to victims that we can only guess at. The general point is that the effects of an intervention may depend on the characteristics of the system in which it is embedded.

An equally fundamental reason why we should hesitate to draw general policy conclusions from these data is that the study offers little insight into why effects associated with arrest emerge. While the authors present their results as a test of the specific deterrence versus labeling theory hypotheses, if specific deterrence is working, we know little about how or why the process occurs. An arrest without prosecution does not, intuitively, seem very serious.[2] Moreover, some data suggest that arrest may work, in part, by breaking up relationships. Even though 86 percent of the arrested men were released within a week, 32 percent of those arrested, but only 10 percent of those separated without arrest, took a week or longer to return to the relationship or never returned at all. Perhaps it

[2] It may of course be that this intuition is wrong (Lempert, 1981-1982), and it may be that an arrest makes the threat of serious punishment for the next offense substantially more credible.

https://doi.org/10.1017/S0023921600017771 Published online by Cambridge University Press

is a good thing for the law to encourage spouse abusers to break off relationships regardless of the woman's wishes, but surely the matter is debatable.

The lack of a good theoretical explanation for why arrests reduce future violence is particularly unfortunate because the small number of cases precludes further meaningful breakdowns of the sample cases. For example, it may be that treatments interact with marital status so that the relative effects of the various treatments on lovers differ from their effects on husbands. Or different kinds of counseling approaches may have been used, and one technique may have been as effective as arrest while others were ineffective or even counterproductive. Also, the broad definition of recidivism used in the study[3] may hide important treatment-related differences. For example, if arrests cause special resentment that leads to extreme violence in a small minority of men while deterring most men from less serious violence, the decision to arrest would, in a minority of cases, have extremely unfortunate consequences that the study would not spot.

Finally, if we take the lack of statistical significance more seriously than perhaps we should,[4] it is possible that policy-makers have drawn precisely the wrong conclusion from this study. When subsequent police contact is the dependent variable, the arrest treatment is significantly more successful than separation but does not differ significantly from counseling. When self-reports are used as the dependent variable, arrest is significantly more successful than counseling but does not differ significantly from separation. Assume the self-reports that are not filtered by the decision to call the police are honest and representative of the entire sample. It would appear from the self-reports that counseling is less effective than arrest but—taking non-significance seriously—that arrest and separation tend to be equally effective. If this is the case, how can the police data be explained? One possibility is that women whose men have been separated but not arrested readily call the police when subsequently victimized, but those whose lovers and husbands have been arrested, although they did not request arrest, are reluctant to do so. Even more reluctant to call the police are those who experienced only

[3] Subsequent property damage and threats of violence as well as actual violence were taken as indicators of recidivism in the interview data, and whatever led the police to be called classified the offender as a recidivist in the police data.

[4] Arrest is clearly the most effective of the three treatments, whether measured by subsequent police contacts or self-reports.

counseling which, as the self-report data suggest, was largely ineffective and may even have been counterproductive. Thus, in the police data differential reporting can explain the entire pattern. And if this explanation is correct, arrest is not only more costly than separation, but it has the unfortunate effect of making victimized women more hesitant to seek police aid in the future.

These criticisms are not really criticisms of the Berk and Sherman study, for as I have said from the standpoint of social science it is an excellent piece of work. These comments do, however, suggest that the results of this research have been prematurely and unduly publicized, and that police departments that have changed their arrest practices in response to this research may have adopted an innovation that does more harm than good.

The example has a number of implications for our discipline. First, we should know more about how the results of certain research come to be widely disseminated. What pressures are there on researchers and granting agencies to publicize the results of their work? Are cautions that accompany careful scholarship inevitably lost in press summaries or interviews? If so, does this say something about how we should or should not publicize what we do? How does the political climate of the time affect the way research is received? Surely, the reception accorded Sherman and Berk's study is in part attributable to the fact that its policy implications were congenial both to law-and-order conservatives and to liberals deeply concerned with women's issues. In short, we need to know more about how research knowledge becomes popular learning so that researchers can intelligently influence the process.

Second, we should remember that the key to generalizing in science is theory. If we understand a process and can explain what is occurring, we have a powerful tool for anticipating what will occur in other situations. If we simply assume that what has occurred in one setting will occur in another, our generalizations will rest on shaky ground whenever the settings differ in important particulars. Making policy on the basis of a single study is always dangerous in part because one study is almost never sufficient to develop reliable theory.

Finally, there is the issue of how much research or how well-developed a theory we should demand before we communicate our results to policy-makers and encourage them to change their behavior to reflect what we have found. After

all, the Sherman and Berk results are theoretically plausible, and if they do generalize an immediate change in how police intervene in minor domestic assaults may save many women from beatings. Here I think we should take a cue from medicine, which painstakingly tests new drugs for safety and effectiveness before putting them into general distribution. It is true that people may and, no doubt, do suffer because the effectiveness of a drug is not sufficiently proved by its first promising test. But they are also saved by careful testing from needless expenditures and from side effects that are worse than the disease, and the credibility of medicine is enhanced in the long run. We should strive for similar testing of law and social science research and theory before we make policy prescriptions to be implemented on a general basis. Along these lines, it is my understanding that the National Institute of Justice, which funded the original Sherman and Berk experiment, is funding a multi-city replication. This is exactly the right response to this potentially important piece of work. Given the low priority too often accorded replications, the NIJ deserves to be congratulated. Widespread publicity and efforts to change law enforcement practices should, however, have awaited the results of the multi-city replication. Instead, they appear to have preceded it.[5]

Turning to the current issue, the last in Volume 18, one finds a series of fine articles. First is Neil Vidmar's piece, "The Small Claims Court: A Reconceptualization of Disputes and an Empirical Investigation." Vidmar's work begins with the perception that people are often not serious about the claims they make when bringing or resisting lawsuits. Plaintiffs often ask for more than they feel they are entitled to, and defendants often deny all liability when they would admit privately that something is owing. Applying this perspective, Vidmar interviewed the parties to small claims courts disputes and found that defendants commonly resisted only a portion of what plaintiffs sought. Thus, it was often the case when a plaintiff received a partial recovery that the defendant had in fact prevailed on most or all of what was actually in dispute. Yet in most small claims court studies a high incidence of partial plaintiff recoveries is interpreted as evidence that

[5] Berk and Newton (1984) have completed a manuscript based on police data which suggest that for certain types of wife batterers—in particular those who in one California county were most vulnerable to arrest—arrest had a preventive effect. The data came from ordinary police reports rather than a field experiment, and in the county studied, unlike Minneapolis, arrest was often followed by other sanctions.

plaintiffs are overwhelmingly successful in small claims litigation. Vidmar's approach, by giving us a better view of what is actually contested, calls this as well as other generally accepted "truths" about the small claims court into question. More generally, Vidmar's study complements the recent literature on sample selection bias by showing that there are fundamental limitations to research that examines court docket information without asking how the docket was generated.

David Engel's article, "The Oven Bird's Song: Insiders, Outsiders, and Personal Injuries in an American Community," is a rare ethnographic glimpse into the legal life of an American rural community. Its focus is on how different segments of the community responded to personal injuries and on the implications of such responses for the likelihood and location of personal injury litigation. Engel's approach differs from Vidmar's, but in an important sense his mission is the same. Like Vidmar he seeks to give us a richer perspective on the social norms and disputes that underlie legal life, and, even more than Vidmar, Engel makes it clear that we cannot understand what is happening in the courts unless we understand the society the court serves and the circumstances in which disputes arise.

The research note that concludes this issue also picks up on the theme that events in court and the statistics that report them may conceal important information or even positively mislead. Dan Lewis, Edward Goetz, Mark Schoenfield, Andrew Gordon, and Eugene Griffin in their article "The Negotiation of Involuntary Civil Commitment" examine how people come to be committed to mental hospitals under new laws that seek to guarantee due process in civil commitment proceedings. They are limited in their ability to generalize by the small number of hospitals and lawyers that they observed, yet their core finding—that voluntary admissions to mental hospitals are often the negotiated outcomes of attempts at involuntary commitment—is so consistent with the literature on plea bargaining in the courts that it would be surprising if the practice were not widespread. If it is, the implications of those studies that reveal declining rates of involuntary mental hospital admissions following the institution of due process protections must be reconsidered.

The other two articles in the issue, "Deterrence and Subjective Probabilities of Arrest: Modeling Individual Decisions to Drink and Drive in Sweden" by Perry Shapiro and Harold Votey, Jr. and "The Impact of Recent Changes in

California Drinking-Driving Laws on Fatal Accident Levels During the First Postintervention Year: An Interrupted Time Series Analysis" by Michael Hilton, deal with the problem of deterrence and the drunk driver. Both articles report a persistent deterrent effect to legal sanctions for drunk driving. These results are important because almost all previous research on legal countermeasures against drunk driving suggests that any deterrent effects that emerge rapidly decay.

Although Hilton and Shapiro and Votey are addressing a similar problem, their approaches are quite different. Shapiro and Votey develop a rational choice model that allows them, in principle, to treat the decision to drink and drive for what it is, individual rather than aggregate behavior. It is the development of this model that constitutes their major achievement, for while their results are interesting and plausible, their data are not sufficiently rich to exploit the full potential of the model.

Hilton, on the other hand, uses familiar ARIMA time series techniques to investigate the early effects of a California countermeasure package. While he finds a marked dip in auto fatalities in the first postintervention year, he is understandably reluctant to attribute this to the countermeasures because the dip is uniform across types of accidents rather than concentrated in categories of accidents (e.g., weekend, nighttime) that are thought to be particularly likely to be alcohol-involved. He does, however, find that serious personal injuries in the alcohol-related categories do disproportionately diminish in the first postintervention year, and some preliminary data from the first nine months of the second postintervention year suggest that fatal accidents have also begun to decline.

I find it difficult to understand how serious personal injury accidents can diminish as a result of alcohol countermeasures without a concomitant diminution in alcohol-related fatalities. Thus, contrary to Hilton, I think even fatal accidents may have been initially deterred. It is at least possible that the observed diminution in what are thought to be the non-alcohol-related categories of fatal accidents results largely from behavioral changes by the subset of drinking drivers who drive at times when drinking is, in the aggregate, less common. If so, California may have discovered a package of reforms that works immediately and persists. Nevertheless, I would not

urge policy-makers in other states to drastically overhaul their laws on the basis of this or any other single study.

Richard Lempert
December 1984

REFERENCES

BERK, Richard A. and Phyllis J. NEWTON (1984) "Does Arrest Deter Wife Battery? An Effort to Replicate the Findings of the Minneapolis Spouse Abuse Experiment." Unpublished manuscript.

KALVEN, Harry, Jr. (1968) "The Quest for the Middle Range: Empirical Inquiry and Legal Policy," in G. Hazard, Jr. (ed.), *Law in a Changing America*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

LEMPERT, Richard O. (1981-1982) "Organizing for Deterrence: Lessons From a Study of Child Support," 16 *Law & Society Review* 513.

SHERMAN, Lawrence W. and Richard A. BERK (1984) "The Specific Deterrent Effects of Arrest for Domestic Assault," 49 *American Sociological Review* 261.