Law and Human Behavior, Vol. 13, No. 3, 1989

Using Psychology to Control Law

From Deceptive Advertising to Criminal Sentencing*

Shari Seidman Diamond†

Legal decision makers often fail to make use of relevant psychological research. In two areas, deceptive advertising and criminal sentencing, legal decision makers *have* welcomed social science research. In each, the research provided has been substantially flawed. Using a commercial that several courts evaluated for deception, I illustrate how the typical study that purports to measure deception produces results that are unnecessarily ambiguous. Then, based on research that looks closely at public responses to criminal cases, I show that the frequently cited survey measures of public preference reflect sentencing preferences for unrepresentative stereotypic criminal offenders. The weaknesses demonstrated in these examples suggest that psychologists can present legal decision makers with a more accurate picture of human perceptions and preferences. If researchers present legal decision makers with informative research when the relevance of research is acknowledged, legal decision makers are likely to become more receptive and more knowledgeable when a new question warrants the application of social science evidence.

Psychologists increasingly appear on the witness stand, in legal briefs, and in legislative hearings to offer their research to the legal system (Melton, 1987). These efforts to use psychological research to influence legal decisions do not always meet with wild enthusiasm or even thoughtful consideration (see, e.g., *Lockhart v. McCree,* 1986).[1] Although researchers on law often conduct research that has no obvious and indeed no expected legal or policy implications, they are

* An earlier version of this paper was delivered as the presidential address for Division 41, the American Psychology—Law Society, at the 1988 American Psychological Association meetings in Atlanta, Georgia. All requests for reprints should be sent to the author at the Department of Psychology, University of Illinois at Chicago, Box 4348, Chicago, IL 60680.

† University of Illinois at Chicago/American Bar Foundation.

[1] 476 U.S. 162 (1986).

0147-7307/89/0900-0239$06.00/0 © 1989 Plenum Publishing Corporation

understandably disappointed when courts reject or ignore crucial empirical research that bears directly on the legal question being decided.[2]

My focus here is on research findings from two areas in which the legal system *has* been responsive to behavioral research, but in which courts have mistakenly relied on research that has been misleading. In both areas, I will argue, behavioral research has presented legal decision makers with an inaccurate picture of the way people understand and evaluate. I will suggest that if psychologists will apply their available methodological tools and theoretical perspectives to these legal questions, they can provide valuable and better data about the behavior of law and a more complete understanding of issues that concern legal actors struggling to reach tough decisions. And because these areas present "trouble cases" for the legal system in which legal doctrine does not provide clear direction, the legal system is likely to be more receptive to the research than when it is asked to evaluate psychological research that raises an apparently settled legal question.[3]

The two areas I will examine are deceptive advertising and public attitudes towards criminal sentencing. These examples represent two strands of work on psychology and law: The first comes from civil law, that vast portion of the law on which psychologists have only recently focused their attention. The second is derived from criminal law, the sector of the law that traditionally captured and still attracts the attention of much research on psychology and law. They both represent areas in which behavioral research has painted an inaccurate picture of how people understand and evaluate. Psychologists can, I think, do better.

DECEPTIVE ADVERTISING

I became interested in deceptive advertising from the vantage point of a psychologist engaged in the practice of law. The relevance of psychology to the regulation of advertising emerged as I encountered case after case that involved consumer perception data. If an advertiser makes a deceptive claim about his or her product, a competitor can bring a civil action in federal court to enjoin that deceptive advertising.[4] Administrative agencies as well as courts have the power to curb false advertising. Both the Federal Trade Commission[5] and the Federal Drug Administration[6] can enjoin deceptive advertisements. The two executives from Beechnut who sold apple juice for infants that did not contain any apples received fines and prison sentences as the result of a prosecution by the Federal Drug Administration for deceptive advertising. Self-regulation also influences the content of national advertising: The television networks each require substantia-

---

[2] For an example of such research, see the articles in *Law and Human Behavior, 8*(1,2) (1984), documenting the conviction proneness of the death-qualified jury.

[3] See, for example, *Barefoot v. Estelle* (1983), in which the Supreme Court rejected the substantial evidence that predictions of violence cannot be made with certainty.

[4] Section 43a of the Lanham Act, 15 U.S.C.A. 1125(a).

[5] Codified at 15 U.S.C. 45.

[6] Codified at 21 U.S.C. 301–392.

tion for any factual claims before they accept a commercial for broadcast,[7] and the National Advertising Division of the Council of Better Business Bureaus (NAD/CBBB)[8] formally evaluates claims of false advertising. The NAD publishes its findings in a monthly newsletter and through negotiation and public dissemination of the newsletter attempts to persuade advertisers to modify or withdraw advertising deemed to be deceptive.

The questions about deceptive advertising would touch psychology very little if explicit falsehood were the only legal basis for a finding of deception. When the makers of Anacin claimed that Anacin can reduce the inflammation that comes with most pain, the express claim concerned the superiority of Anacin over Tylenol for reducing inflammation (*American Home Products Corp. v. Johnson & Johnson,* 1978).[9] A dictionary would be the appropriate reference for the accepted meaning of, for example, the word *inflammation;* a psychologist would have little to say.

But the legal standard for deception is *not* simply literal falsehood. It includes instead the way consumers understand the commercial message. As a result, the legal test of deceptive advertising is fundamentally a behavioral test—a psychological test of perception. Courts have recognized that "a statement acknowledged to be literally true and grammatically correct [may] nevertheless [have] a tendency to mislead, confuse or deceive" (*American Brands, Inc. v. R. J. Reynolds Co.,* 1976, p. 1357).[10] This tendency to mislead cannot be judged without considering how a communication is perceived and "the court's reaction is at best not determinative and at worst irrelevant. The question in such cases is—what does the person to whom the advertisement is addressed find to be the message?" (*American Brands, Inc. v. R. J. Reynolds,* 1976, p. 1357). The accuracy of what consumers perceive the commercial message to be is thus the standard against which deceptive advertising is evaluated. The makers of Tylenol were able to stop Anacin's claims about reducing the "inflammation that comes with most pain" although the claim was literally true. They did it by convincing the court that consumers understood the ad to imply that Anacin offered greater general pain relief than Tylenol (*American Home Products Corp. v. Johnson & Johnson,* 1978, p. 167).[11]

Parties bringing or defending claims of deceptive advertising have responded to the behavioral standard announced by the courts by supplying the courts with consumer perception studies that purport to measure the tendency of a particular commercial to mislead the viewing public. In view of the legal standard, it may not be surprising that this research is regularly accepted and apparently relied on by

---

[7] See, e.g., the Capital Cities/American Broadcasting Company, Inc., Department of Broadcast Standards and Practices, "Advertising Standards and Guidelines" (1986).

[8] The NAD decided 103 cases in 1988 (NAD/CBBB, 1989). In 75 cases the advertisers modified or discontinued the allegedly deceptive advertising, in 27 cases the NAD concluded that the claim was substantiated, and one case was referred to the National Advertising Review Board for a further determination.

[9] 577 F.2d 160 (2d Cir. 1978).

[10] 413 F. Supp. 1352 (S.D.N.Y. 1976).

[11] 577 F.2d 160 (2d Cir. 1978).

the courts. What is distressing is that in this arena where behavioral science research is willingly received by the courts, the caliber of the research is so woefully inadequate to the task.

In the typical survey of consumer perception, the marketing researcher brings members of the public who regularly make purchases in the product category into a survey facility and shows them the allegedly misleading commercial. If it is a television commercial, respondents usually view the commercial several times. The commercial is usually the only stimulus they view. After viewing it, respondents are asked, "What were the main points that the commercial was trying to get across?" or some other open-ended question that asks the respondent to recall the messages contained in the commercial. Interviewers probe for further recall of the commercial's content. The verbatim responses to these open-ended questions are analyzed for evidence of deceptive messages conveyed to the respondents. Follow-up questions are sometimes added in which respondents are asked about specific statements or topics introduced in the commercial. For example, in *Johnson & Johnson v. Carter-Wallace* (1981)[12] the makers of Johnson's baby oil alleged that a commercial for Nair with baby oil falsely conveyed the message that the product independently softened and/or moisturized the user's skin. Respondents in the consumer perception test were first asked, "What ideas did the advertiser try and get across about Nair in their commercial?" A follow-up question accepted by the Court was, "What if anything is good about having baby oil in the product?"

Multiple-choice follow-up questions are also accepted if they are not perceived as leading, but some courts have become reasonably sophisticated in recognizing the potential biases in some question structures. For example, physicians who viewed an ad for Tylenol were asked whether the commercial indicated that the ingredients in its two major competitors, aspirin and ibuprofen, caused health hazards that were similar. They were asked to indicate whether, when taken at nonprescription dosage levels, aspirin and ibuprofen have the same likelihood of causing gastric ulcers, about the same likelihood, somewhat different likelihoods, very different likelihoods, or none of the above (*American Home Products Corp. v. Johnson & Johnson,* 1987, p. 581).[13] In fact, the commercial said that both aspirin and ibuprofen can cause gastric irritation, but said nothing about their tendency to cause ulcers. The court acknowledged that pure guessing would produce 20% who endorsed each of the five possible choices, noted that "somewhat different" and "very different" were so obviously false that they were not genuine choices and recognized that the negative wording of the choice, "none of the above," decreased the likelihood that it would be chosen by respondents.

Not all courts are so careful, however. In another of the pain-reliever cases, the makers of Tylenol charged that a commercial for Anacin deceptively implied that Maximum Strength Anacin was a stronger analgesic than Extra-Strength Tylenol (*McNeilab, Inc. v. American Home Products Corp.,* 1980).[14] A bar graph

---

[12] 487 F. Supp. 740 (S.D.N.Y. 1981).
[13] 654 F. Supp. 568 (S.D.N.Y. 1987).
[14] 501 F. Supp. 517 (S.D.N.Y. 1980).

in the commercial showed the comparative amount of pain reliever in "regular strength products," "extra-strength products," and Maximum Strength Anacin. In the follow-up questions of a consumer perception test, respondents were asked, "From what you have seen or heard in the commercial, do you think that they were attempting to compare Maximum Strength Anacin to the brand of pain reliever you, yourself, use most often?" The court rejected expert testimony by the defense that the question directed respondents towards a yes answer.

Finally, the quantitative results from the consumer perception survey are evaluated to determine whether they establish that the advertisement *tends* to mislead. Although no single standard has been adopted for this tendency, one influential treatise (McCarthy, 1984, Section 32:54) has suggested using the percentage figure accepted as evidence of likelihood of confusion in trademark surveys—a percentage that can be as low as 7% (*Wendy's International, Inc. v. Big Bite, Inc.*, 1980).[15] Courts hearing deceptive advertising cases have found survey figures to be sufficient when they indicated that "over 25%" of consumers received a misleading message from the ad (*R. J. Reynolds Tobacco Co. v. Loew's Theatres, Inc.*, 1980).[16]

Much of the testimony and argument in deceptive advertising cases address the adequacy of the consumer perception research. Sometimes both sides produce surveys; at other times, the defense relies on criticisms of the plaintiff's survey. Courts have addressed several general questions about the methodology of such research. First, they have considered the selection of respondents. Respondents in consumer perception tests are typically convenience samples of 200–300 consumers who have recently made purchases in the product category. Courts accept the mall intercept samples typically used, recognizing that it would be difficult to show an allegedly deceptive commercial to a random sample of consumers and relying on testimony that large companies use such samples in making major marketing decisions about their product.

Second, courts have rejected the day-after recall test in which the viewer is questioned on his or her reaction to commercials that aired on the previous day. They instead favor a forced exposure approach in which consumers view the commercial in a survey facility (*American Home Products Corp. v. Johnson & Johnson*, 1977).[17] The court in the case that set this standard formed its opinion based on the verbatim responses to a day-after recall survey. Because many of the responses were rather confused, the court concluded that the forced exposure approach would more accurately reflect what the average consumer who heard and saw the ad took the message(s) to be.

The reasoning seems faulty, but is widely accepted. If consumers are to be protected against misimpressions that they get from viewing commercials, it seems appropriate to measure the impressions they get when they view the commercials as they would ordinarily be exposed to them outside the test. The court in *American Home Products* assumed that the forced exposure approach would

---

[15] 596 F. Supp. 816 (S.D. Ohio 1980).
[16] 511 F. Supp. 867 (S.D.N.Y. 1980).
[17] 436 F. Supp. 785 (S.D.N.Y. 1977), *affirmed* 577 F.2d 160 (2d Cir. 1978).

be most sensitive to deception. In fact, a commercial may be less likely to deceive viewers who watch it in a forced exposure test because they are primed to be attentive to subtle qualifications and disclaimers.

The third methodological point that has received the most attention is the design of the questions. Judges, like all lawyers, see themselves as wordsmiths. Concerned that closed-ended questions may be necessarily leading, the courts have expressed a strong preference for open-ended questions (e.g., What were the main ideas presented in this commercial?). As a result, much courtroom argument in these cases has focused on the interpretation of the verbatim responses to these open-ended questions. In one case where the plaintiff's and defendant's experts strenuously disagreed on the number of allegedly deceptive responses—the plaintiff arguing that 43% were deceptive—the court did its own analysis and identified no more than 15% that indicated deception, (*Coca-Cola Co. v. Tropicana Products, Inc.*, 1982).[18]

At the same time, some courts have accepted that tests of free recall can provide an incomplete measure of the message that respondents carry away from a communication. They have even complained when direct questions on a crucial point were *not* asked. The makers of Nuprin claimed, "Two little Nuprin are stronger than Extra-Strength Tylenol" (*McNeilab v. Bristol-Myers,* 1986).[19] The question was whether consumers perceived that Nuprin's superiority over Extra-Strength Tylenol was achieved only when *two* Nuprin tablets were taken. The makers of Tylenol asked consumers several open-ended questions to assess what the commercial was trying to get across. Few respondents mentioned the dosage of Nuprin in the open-ended questions, and plaintiff's expert testified that consumers did not understand that the superiority claim was related to the two-tablet dosage. The court criticized the survey as "designed to avoid a direct consumer response" on the issue in the case because the survey did not ask, "How many tablets of each product were compared?" (p. 91).

Extensive methodological discussions fill the pages of court decisions on deceptive advertising, and judges often bemoan their inability to evaluate the evidence. They neglect, however, a crucial change in the standard methodology, one familiar to all well-trained psychologists, that would avoid the ambiguous results most judges face in evaluating marketing surveys that purport to test consumer deception.

The question in a case of allegedly deceptive advertising is: Does the advertisement tend to deceive or mislead the viewing public? The question is causal—are perceptions altered by the communication? Yet nearly all studies in published federal court opinions that purport to show the impact of deceptive advertising on consumers are surveys in which a single group of respondents is questioned on their impressions of the allegedly deceptive commercial. Using this single group posttest-only design (Cook & Campbell, 1979), it is impossible to evaluate whether the responses are solely the product of the advertisement, or whether

[18] 538 F. Supp. 1091 (S.D.N.Y. 1982), *reversed on other grounds*. 690 F.2d 312 (2d Cir.).
[19] 656 F. Supp. 88 (E.D. Pa. 1986).

they are the result of preexisting impressions of the product being advertised, expectations about what commercials are likely to say, or responses based solely on the way that the questions in the survey are asked.[20]

A pain reliever is advertised. Consumers expect the ad to extol the virtues of the advertised product for pain relief. When asked for ideas that the ad conveys or asked, "Based only on what the commercial said, would Maximum Strength Anacin contain more pain reliever, the same amount of pain reliever or less pain reliever than the brand you, yourself, currently use most often?" a number of the respondents naturally guess that the advertisement claims the superiority of the advertised product. Yet the ad may have included *no* claim of superiority and may have had *no* effect on the perceptions of the consumers who viewed it. The causal impact of the ad on perceptions was not tested in the survey.

It could be. The causal question is testable by traditional experimental methods. To demonstrate one way it would be used in a case of deceptive advertising, Kwang Park and I tested perceptions of a Tropicana orange juice commercial that was the subject of extensive litigation. The District Court struggled with the results of a survey, finally decided that 15% of the consumers tested were deceived, and concluded that 15% was not enough. The Court of Appeals, in contrast, thought that 15% was not insignificant.[21]

The television commercial for Tropicana showed Bruce Jenner squeezing the juice from an orange into a pitcher. Then, pouring the juice directly into the Tropicana carton, he said, "Pure, pasteurized juice as it comes from the orange." Subsequently he described Tropicana as "the only leading brand not made with concentrate and water." Coca-Cola, the maker of Minute Maid orange juice, charged that the advertisement inaccurately conveyed the idea that Tropicana is fresh squeezed, unprocessed orange juice. The advertisement is literally ambiguous: Juice that is pasteurized cannot be unprocessed and fresh squeezed into the carton.

We tested the effect of the allegedly deceptive portions of the ad, the visual picture of Jenner pouring the juice into the carton and the verbal description "as it comes from the orange," by measuring respondents' perceptions of the ad with and without those portions present. Respondents were randomly assigned to one of three groups. The first group saw the original ad; the second listened to an audio tape of it; the third group listened to an audio tape without the phrase "as it comes from the orange." All respondents were asked a series of open-ended questions to get their ideas on what message the ad was trying to convey. Half the respondents were then given a series of multiple-choice questions on the messages

---

[20] In their research on miscomprehension of TV communications, Jacoby and Hoyer (1982) found that respondents gave incorrect answers to an average of 28% of the true–false test items used to assess viewer comprehension of nondeceptive commercial advertising. While this figure may be somewhat inflated (e.g., because respondents were not told they could answer "don't know," Ford & Yalch, 1982), the results indicate the potential magnitude of the background noise that will be measured even in nondeceptive advertising.

[21] The Court of Appeals also found that the commercial was literally false and enjoined the commercial on that basis, *Coca-Cola Co. v. Tropicana Products, Inc.* (1982).

in the commercial; the other half answered a set of true–false questions. The results demonstrated that deceptive responses rose significantly on all three types of measures from the purged audio commercial to the full visual commercial. The only plausible explanation for the difference was the difference in what the respondents were exposed to—the content of the commercial.

Adopting an experimental paradigm for the analysis of deception also reduces concerns that the wording of a question is leading: It should be equally leading to respondents who view the control version of the ad. Similarly, if open-ended questions are used, and the responses are analyzed according to the same standards in the experimental and control groups, the differences can only be attributed to the allegedly deceptive content of the ad. Courts may, of course, decide that mere statistical significance is not sufficient to warrant injunctive relief and may demand a showing that the ad affects perceptions in a specified amount, but use of the experimental method would at least make it easier for courts to detect a genuinely deceptive effect. It would also systematically avoid a spurious finding of deception.

The tools of the experimental method are readily available for use in the study of deceptive advertising. The next step is simply the application of those tools in court by an adventurous litigant either spurred on by a demanding court or stimulated by a persistent expert.[22] That is not to underestimate the difficulty of persuading a litigant or a court to try a new approach. It simply indicates that the path is clear: The application of a familiar methodological device—the control group—to a legal question that is already defined as behavioral science terrain.

CRIMINAL SENTENCING

My second example of a sociolegal question ripe for psychological attention can profit from an infusion of both the method and the theory of psychology. Until recently, the primary official goals for criminal sentencing were rehabilitation and deterrence (see, e.g., Model Penal Code, 1962). These instrumental effects of punishment received attention from legislators, judges, and correctional officials. Then came a disillusionment with the effectiveness of rehabilitation programs (Lipton, Martinson, & Wilks, 1975), and evidence appeared that the deterrent effects of punishment were small (Blumstein, Cohen, & Nagin, 1978). This was accompanied by a shift in philosophy toward a *just deserts* model of punishment. According to the just deserts model, offenders were to be given the punishment they deserved in light of the offense they committed—no more and no less, and unaffected by an assessment of whether an appropriate sentence could reform them (see, e.g., von Hirsch, 1976). Fixed sentences and determinate sentencing

[22] Control groups have been used in trademark cases where the question is whether consumers believe that a product is produced or sponsored by the holder of a valid trademark. They have also been proposed for FDA evaluations (Jacoby & Small, 1975). It is unclear why they have been neglected in federal court decisions on deceptive advertising.

schemes that specified the character and length of the sentence replaced the once standard indeterminate sentences that would be adjusted based on the offender's postsentencing progress.

One difficulty with this shift in ideology was that the standards for a good or appropriate sentence changed. No longer could a sentence be evaluated according to whether it reduced recidivism or parole violation rates. Now the standard became the perceived fairness of the severity of the punishment.

Whereas professional "treatment specialists" were appropriate judges of the reform capabilities of a sentence, the views of the general public became increasingly relevant for a sentencing policy based on perceived fairness (Blumstein & Cohen, 1980). Enter the public opinion polls.

In the push to replace indeterminate sentencing laws with fixed and mandatory penalties, legislatures have listened to the echo of their constituents' demands in the public opinion polls. In Illinois, for example, legislators in 1981 passed a bill that makes residential burglary a Class One felony subject to a mandatory four-year prison term.[23]

What the public opinion polls purport to show is a public preoccupied with the problem of crime, critical of courts that are too lenient, and in favor of tougher penalites for convicted offenders. For example, 85% of respondents to a 1986 public opinion poll reported that their local courts were not dealing harshly enough with criminals (Jamieson & Flanagan, 1987).

The results seem clear, and the pattern of findings varies little from survey to survey—a fearful, punitive public wants courts to respond with harsher penalties. The courts, lagging behind public opinion, need the constraints of legislative action to make them meet public standards.

Yet there is reason to question the public punitiveness that the polls appear to demonstrate. Are the sentences that the public demands really different from those that judges give out? After all, most citizens surveyed in the public opinion polls have little direct contact with the criminal courts. They obtain most of their information indirectly—from the newspapers, television, and the stories of relatives and friends. If the reports are not representative, and much research suggests they are not (e.g., Graber, 1980), the picture of judicial performance obtained from these sources should be distorted.

The most direct way to test whether the public does indeed favor more severe sentences than does the judiciary is to compare the sentences the two sets of respondents would give to the same set of offenders. I got the opportunity to develop such a test when I was invited to participate in a series of judicial seminars on sentencing for Illinois judges. Over 100 Illinois judges participated in the seminars. As part of their participation they sentenced four offenders after reading their presentence reports and viewing videotapes of their sentencing hearings. Later, with the help of Judge Warren Wolfson, Loretta Stalans and I presented the same materials to a sample of Cook County jurors who sentenced the same four offenders (Diamond & Stalans, 1989). The cases included a burglary by a repeat

---

[23] Ill. Rev. Stat. 1987, ch. 38, par. 19–3.

offender with a drug addiction, a sale of cocaine, a purse snatch from an elderly victim who was seriously injured in the course of the robbery, and an assault with a knife in the course of a barroom fight. We selected these cases for the judicial seminars because they were standard criminal offenses that were possible, but not definite, candidates for prison sentences.

The results were surprising. On none of the cases did the jurors give more severe sentences than the judges gave. In fact, the judges were significantly more severe than were the jurors. And this occurred despite the fact that most jurors we tested thought that courts were too lenient. On each case we asked the jurors if they thought their own sentences were more severe, the same, or more lenient than those of the average Illinois judge. Even those who thought they were more severe than the average judge in fact gave sentences that were no more severe than the mean judicial sentence. It appears that at least part of the public's apparent discontent with judicial action is based on an inaccurate picture of judicial sentencing behavior.

But there is a second explanation that comes from psychological theory and research on decision making. Who is the public thinking about when it advocates more stringent penalties? Doob and Roberts (1984) asked respondents who they had in mind when they were asked to evaluate whether courts were too harsh, too lenient, or about right. Respondents who said that the courts were too lenient in sentencing offenders were more likely to report they were thinking about violent or repeat offenders than were respondents who indicated that the courts were about right. This study gives a clue that the availability heuristic is operating when the public responds to global questions about the treatment of offenders. Tversky and Kahneman (1974) have shown that we select from our repertoire of experiences the ones that are easiest to recall, and that when we generalize based on these examples, our judgments may be biased by the disproportionate impact of available instances. Moreover, vividness appears to contribute to availability of information (Nisbett & Ross, 1980; Reyes, Thompson, & Bower, 1980; but see, Taylor & Thompson, 1982). Thus, our image of the typical burglary offender is not the accurate picture of a young, unarmed offender who slips in and out when no one is around, but rather of the professional burglar, armed with a weapon, who confronts the startled homeowner or ransacks the premises. This biased construction of reality can cause respondents to base their sentencing opinions on images of offenses and offenders that are disproportionately severe.

Other research by Doob and Roberts (1984) shows what can happen to general support for harsh penalties when respondents cannot fill in the picture of the offender with demon characteristics. They varied the amount of information respondents received about a criminal case before sentencing the offender. Half of the subjects read a newspaper account of the case, and the remaining subjects read the court transcript. Of the respondents who read the newspaper account, 80% indicated that the judge was too lenient, while only 14.8% of those who read the court transcript indicated that the judge was too lenient.

Similarly, Ellsworth (1978) found that respondents were less punitive when they were asked to make concrete sentencing judgments than when they were

asked about abstract sentencing preferences. When asked, "For the crime of X (killing a policeman, beating a wife to death, premeditated murder), do you believe in capital punishment or are you opposed to it?," 50%–66% of the respondents she surveyed said they were in favor of it. Respondents were also asked if they favored capital punishment in a series of concrete cases in which the offender and the circumstances of the offense were described. In no concrete version of any of the offenses, however damaging, did the number of respondents favoring the death penalty exceed 34%.

The impact of the specificity of the case on judgments is an explanation that fits these results, those of Doob and Roberts, and the apparent leniency of the jurors in our own research. The general survey question evokes the image of an offense and offender that is far more extreme than the actual profile of the average offender found guilty of that offense. Once the human details of an offense and offender are described, the average offender appears far less deviant, powerful, and dangerous; severe punishment appears less justified.

Does this body of research have any policy implications? At the least, it suggests that a better-informed public would find judicial behavior more acceptable and argues for an attempt to educate officials to the softness of poll data. At the most grandiose, it suggests a drastic remedy for public dissatisfaction with sentencing. If the public is not a vengeful, punitive dragon, and the task of sentencing is to determine the just level of punishment for a particular offender in light of his offense, then the argument for a professional sentencing judge loses much of its force. Moreover, if public dissatisfaction with the courts is fed by inaccurate perceptions of court actions, then a way to reduce dissatisfaction is to increase contact between the public and the sentencing decision.

De Tocqueville (1876), who was an enthusiastic supporter of the jury, claimed that the primary virtue of the jury system was its educative function in informing the lay public about the legal system. Currently, in most states and in the federal system, the jury is dismissed once it decides to convict or acquit a criminal defendant—unless the case is a capital one in which the jurors will decide whether the offender will be sentenced to death. Why not extend the jury's role to the sentencing arena more generally?

Judges complain that sentencing is their hardest task (Gaylin, 1974). The legal rules that confine sentencing decisions are surely less complex than those that govern judgments about guilt, and desert is clearly a normative judgment. A possible argument against using an ad hoc tribunal like a jury for sentencing is that it will promote sentence disparity. Every study of judicial sentencing reveals a significant variation among judges in the sentences they would give identical offenses and offenders (see, e.g., Diamond & Zeisel, 1975). Arguably, an ad hoc decision maker like a jury would only exacerbate the usual problems of sentencing disparity. But note that one major benefit of a decision from a group like a jury rather than an individual decision maker like a judge is that the group decision pools individual preferences and is thus less likely to be the product of a single, extreme decision maker.

Moreover, recent attempts to reduce sentence disparity can be used to shape

the jury's sentence determination.[24] A number of jurisdictions have instituted sentencing guidelines that structure the task of sentencing (e.g., Minnesota, Pennsylvania). Further, if the technical aspects of sentencing are troublesome, or there is need for more direct input on the "going rate" for offenses of this type, perhaps the answer is a mixed tribunal on which the judge joins the jury to decide on the sentence.[25] The mixed tribunal is the forum used for sentencing in Austria and West Germany. West Germany uses a mixed tribunal of lay and professional judges to decide both whether to convict and what sentence to give a convicted offender. A study of the decision making on these mixed tribunals showed that the lay judges were far more active and influential on the sentencing decision than in deciding whether to convict (Casper & Zeisel, 1972). It is just possible that use of sentencing jurors in U.S. courts could both legitimize the sentencing decision and make it better reflect public conceptions of desert.

CONCLUSIONS

We have now examined one research arena in which psychological research can improve the methods used to reach legal decisions and another in which a more psychologically informed view of public demand from the legal system might ease the conflict between court and public. The special tools and theories of psychology are not always appreciated—as the current Supreme Court has most emphatically demonstrated. But I would suggest at least one approach to the legal system appropriate for researchers who want ultimately to influence legal policy: Be more responsive to the trouble cases of the legal system when behavioral science is identified as a source of knowledge. When the researcher begins with a legal forum that believes social science is likely to have some answers, the probability of a receptive ear is increased. Thus, on topics from the child and family relationship to how consumers typically use products and interpret labels, legal sources have invited social inquiry and expertise. If some of these topics do not have the pathos of racial discrimination or unfair imposition of the death penalty, they lay the groundwork for judicial and legislative understanding and respect for psychological research. If psychologists do not ensure that the quality of the social science research presented to courts is high, we risk not only bad legal decisions on these topics, but also a reduced willingness to seriously consider social science data in other areas. An exasperated court in a recent deceptive advertising case rejected the surveys produced by both sides. The court said:

> It is difficult to believe that it was a mere coincidence that when each party retained a supposedly independent and objective survey organization, it ended up with survey questions which were virtually certain to produce the particular results it sought. This strongly suggests that those who drafted the survey questions were more likely knaves than fools. If they were indeed the former, they must have assumed that judges are the latter. *American Home Products Corp. v. Johnson & Johnson* (1987, p. 582).[26]

---

[24] For a description and evaluation of some of these efforts, see Cohen and Tonry (1983).

[25] Langbein (1981) has suggested that a mixed court composed of laypersons and judges could improve methods of obtaining convictions in the U.S. criminal justice system.

[26] 654 F. Supp. 568 (S.D.N.Y. 1987).

This was a court willing to hear social science data and disappointed with the result. It would hardly be surprising to find the court less receptive to and more suspicious of the next piece of social science evidence placed before it. It is up to us to see that the judiciary and other actors in the legal system come to understand the meaning and policy implications of social scientific research. If we do not work to educate legal decision makers so they can distinguish between good research and bad, we cannot complain when courts or legislatures give credence to misleading data or reject our own impeccable research findings.

Cases

American Brands. Inc. v. R. J. Reynolds Co., 413 F. Supp. 1352 (S.D.N.Y. 1976).
American Home Products Corp. v. Johnson & Johnson, 436 F. Supp. 785 (S.D.N.Y. 1977).
American Home Products Corp. v. Johnson & Johnson, 577 F.2d 160 (2d Cir. 1978).
American Home Products Corp. v. Johnson & Johnson, 654 F. Supp. 568 (S.D.N.Y. 1987).
Barefoot v. Estelle, 463 U.S. 880 (1983).
Coca-Cola Co. v. Tropicana Products, Inc., 538 F. Supp. 1091 (S.D.N.Y. 1982).
Coca-Cola Co. v. Tropicana Products, Inc., 690 F.2d 312 (2d Cir. 1982).
Johnson & Johnson v. Carter-Wallace, 487 F. Supp. 740 (S.D.N.Y. 1981).
Lockhart v. McCree, 476 U.S. 162 (1986).
McNeilab v. Bristol-Myers, 656 F. Supp. 88 (E.D. Pa. 1986).
McNeilab Inc. v. American Home Products Corp., 501 F. Supp. 517 (S.D.N.Y. 1980).
R. J. Reynolds Tobacco Co. v. Loew's Theatres, Inc., 511 F. Supp. 867 (S.D.N.Y. 1980).
Wendy's International, Inc. v. Big Bite, Inc., 596 F. Supp. 816 (S.D. Ohio 1980).

REFERENCES

American Law Institute (1962). *Model Penal Code.*

Blumstein, A., & Cohen, J. (1980). Sentencing of convicted offenders: An analysis of the public's view. *Law & Society Review, 14,* 223.

Blumstein, A., Cohen, J., & Nagin, D. (Eds.). (1978). *Deterrence and incapacitation: Estimating the effects of criminal sanctions on crime rates.* Washington, DC: National Academy of Sciences.

Casper, G., & Zeisel, H. (1972). Lay judges in the German criminal courts. *Journal of Legal Studies, 1,* 135.

Cohen, J., & Tonry, M. (1983). Sentencing reforms and their impacts. in A. Blumstein, J. Cohen, S. Martin, & M. Tonry (Eds.), *Research on Sentencing: the Search for Reform.* Washington, DC: National Academy of Sciences.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally.

De Tocqueville, A. (1876). *Democracy in America.* New York: Doubleday.

Diamond, S. S., & Stalans, L. J. (1979). The myth of judicial leniency in sentencing. *Behavioral Sciences and the Law, 7,* 73.

Diamond, S. S., & Zeisel, H. (1975). Sentencing councils: A study of sentence disparity and it reduction. *University of Chicago Law Review, 43,* 109.

Doob, A. N., & Roberts, J. V. (1984). Social psychology, social attitudes and attitudes toward sentencing. *Canadian Journal of Behavioral Science Review, 16,* 269.

Ellsworth, P. (1978). Attitudes toward capital punishment: From applications to theory. Paper presented at meetings of Society for Experimental Social Psychology, Princeton, New Jersey.

Ford, G. T., & Yalch, R. (1982). Viewer miscomprehension of televised communication—A comment. *Journal of Marketing, 46,* 27.

Gaylin, W. (1974). *Partial justice*. New York: Knopf.

Graber, D. (1980). *Crime news and the public*. New York: Praeger.

Haney, C. (Ed.). (1984). Special Issue of *Law and Human Behavior, 8* (whole issue).

Jacoby, J., & Hoyer, W. D. (1982). Viewer miscomprehension of televised communication: Selected findings. *Journal of Marketing, 46,* 12.

Jacoby, J., & Small, C. (1975). The FDA approach to defining misleading advertising. *Journal of Marketing, 39,* 65.

Jamieson, K. M., & Flanagan, T. J. (1987). *Sourcebook of Criminal Justice Statistics—1986*. Washington, DC: U.S. Government Printing Office.

Langbein, J. (1981). Mixed court and jury court: Could the Continental alternative fill the American need? *American Bar Foundation Research Journal, 1981,* 195.

Lipton, D., Martinson, R., & Wilks, J. (1975). *The effectiveness of correctional treatment: A survey of treatment evaluation studies*. New York: Praeger.

McCarthy, T. (1984). *Trademarks and unfair competition* (2d ed.).

Melton, G. (1987). Bringing psychology to the legal system: Opportunities, obstacles, and efficacy. *American Psychologist, 42,* 488.

National Advertising Division/Council of Better Business Bureaus, Inc. (1989). *NAD Case Report,* January 16.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Reyes, R. M., Thompson, W. C., & Bower, G. H. (1980). Judgmental biases resulting from differing availabilities of arguments. *Journal of Personality and Social Psychology, 39,* 2.

Taylor, S. E., & Thompson, S. C. (1982). Stalking the elusive "vividness" effect. *Psychological Review, 89,* 155.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124.

von Hirsch, A. (1976). *Doing justice: the choice of punishment*. New York: Hill & Wang.