

SURVEY DATA AS EVIDENCE VS. SURVEY DATA FOR DECISION

Irving Crespi, Irving Crespi and Associates

The use of survey data as legal evidence involves considerations that differ substantially from those that arise in other applications, such as marketing research. Failure to understand how the working of the legal system affects the admissibility and evidentiary value of survey data can lead to frustration and irritation on the researcher's part. It can also lead to a court's rejection of survey results that a marketing manager might consider very useful. By describing how surveys are used in litigation and contrasting this with their use in science and for making marketing decisions, this paper seeks to clarify what makes for the effective use of survey data as evidence.

The use of survey data as evidence is now common. Yet, until relatively recently most courts applied the hearsay rule to exclude survey data. Even now, when survey data are commonly accepted by courts, statistical information of all kinds is often treated by courts as having limited value compared with anecdotal reports that social scientists would consider of lesser value. In 1981, for example, a court commented in deciding a discrimination case that:

We find it very damaging to plaintiffs' position the fact that not only was their statistical evidence insufficient, but that they failed completely to come forward with any direct or anecdotal evidence of discriminatory employment practices by defendants. (*Garcia v. Rush Presbyterian-St. Luke's Medical Center*, 660 F. 2d 1217)

The Supreme Court has expressed a similar attitude toward statistical data, describing their value as contingent upon the uniqueness of individual circumstance:

...statistics are not irrefutable. They come in infinite variety and, like any other kind of evidence, they may be rebutted. In short, their usefulness depends on all of the surrounding facts and circumstances. (*International Brotherhood of Teamsters v. United States*)

As might be expected, social scientists bridled at the earlier exclusion of survey data as hearsay, and they continue to resent the secondary status still given by many courts to data gathered by one of the more valuable methodological tools of their discipline. Nonetheless, there are substantive reasons underlying legal barriers to the use of survey data. Although the hearsay barrier to admissibility is no longer a serious problem, understanding why it once was can help explain why the evidentiary value of surveys may still be subject to question. To this purpose, let us contrast the use of surveys in social science with their use in litigation and in marketing.

Scientific, marketing, and legal "truths"

Social science, marketers, and the American legal system are each concerned with establishing "truth," but they operate under very different sets of definitions, precepts and procedures for achieving that goal. Functioning as a social scientist, the survey researcher seeks to document and cast light on social phenomena, with the development of theories that adequately explain social behavior as his ultimate goal. "Truth," in this

context, is not absolute; it is verifiable, general explanations of social behavior that are subject to continuous revision as new knowledge is obtained. Survey results represent a major class of data used to test those explanations. Furthermore, a fundamental criterion for accepting a scientific report is that peers evaluate the adequacy of the methods used to collect data and to analyze them and, as the final test, that peers replicate the study.

Marketers and courts both seek valid information about the world upon which they can base decisions. Both have turned to survey research as a scientifically established way of getting such information. But the ways in which they apply survey data, and the criteria they use for accepting the data, differ in significant ways. Both, of course, are concerned that proper methodology be used -- that sample designs, questionnaires, interviewing methods, and analytical procedures be appropriate to the problem at hand. But, once the methodological worth of a survey is established, marketing managers are far less restrictive than are courts regarding the "admissibility" of a report.

The marketing manager has to decide what future course of action to take, and seeks whatever information he can get that will help him make that decision. He may accept partial information, he may accept inferences that go beyond the data, he may examine reports bearing on at best analogous problems -- anything that he feels gives him a better understanding of the nature of his problem and the likely outcomes of alternative courses of action. When nothing better is practical, he will knowingly accept survey data based on methods that he knows to be less than state of the art. Thus, budget limitations and time pressures will lead even the most sophisticated to accept non-probability sample designs, low completion rates, and less than ideal measuring instruments. And even though he is normally under considerable time pressure to reach a decision, the marketing manager is also under pressure to change that decision if the desired outcomes are not achieved.

The more information the better is the marketing manager's operating principle, so long as the information has action implications, and so long as he has grounds for believing that it is not wrong. Personal judgment as to how actionable the data are, rather than institutionalized procedures, govern how much weight he gives to any one survey. The one type of survey that a marketing manager is sure to ignore, no matter how valid it may be, is the "interesting" survey that does not help him make a decision. Personal biases and private agendas can, and do, affect the decisions of marketing managers, but exclusion of survey data is not based on binding external principles that, from a survey researcher's perspective, may seem to be arbitrary and irrelevant.

The "truth" that a court seeks to establish contrasts with the marketing manager's truth. Courts want to ascertain the facts of an individual case that has already occurred. Trials are not held to establish a general principle explaining propensity to act in certain ways. They are held to decide which of conflicting assertions accura-

tely describes what occurred in a specific situation. Verdicts are supposed to be reached by weighing all the evidence that has been presented to see if a sense of conviction can be reached as to what really happened. Merely believing that one suspects he has a fairly good idea as to the facts of the case is not sufficient.

Subject to appeal, once a verdict is reached, it is binding on all parties. Regardless of how well they describe the world generally, scientific reports, including those based on statistical surveys, are considered legally relevant only if they cast light on the truth of conflicting contentions concerning the specific event, or set of events, under litigation. This is why anecdotal reports that specify "surrounding facts and circumstances" are highly valued as evidence. And this is also why survey data are typically considered only in conjunction with anecdotal and direct testimony, rather than being relied on as the primary basis for reaching a verdict.

The hearsay rule and the adversary system

Instead of relying on a presumably objective method for gathering and analyzing data to establish the truth of the case at issue, American courts require the confrontation of testimony from the contending litigants -- the adversary system. Among other things, that system requires that individuals testify as to what they, themselves, know directly; and not about what someone has told them. Underlying this requirement is the principle that by subjecting the testimony of various witnesses as to what they know (or claim to know) to cross-examination, the credibility of conflicting testimony can be evaluated by reasonable persons, that is, by jurors and judges. Hearsay -- what someone was told by someone else as opposed to what was directly observed -- is consequently excluded as evidence.

Surveys that rely on interviews were automatically suspect as hearsay since they involve reports on what interviewers are told by someone else and not on their direct knowledge. The traditional legal position was that only if respondents to a survey were subject to cross-examination could its results be admissible as evidence. The sheer impracticality of cross-examining more than a token number of respondents, plus the insistence of survey organizations that the anonymity of respondents be protected, made even the best of surveys subject to exclusion under the hearsay rule.

Overcoming the hearsay barrier required redefining the nature of survey interviews and results. Rather than treating survey results as a compilation of testimony from a large number of individuals, they are now introduced as data about, for example, the state of mind of a defined population. Once this approach is accepted, the persons whose credibility is at issue, and who therefore should be subject to cross-examination, are those who conduct the survey. If a court so decides, they might include interviewers as well as professional staff.

Courts have repeatedly ruled that expert testimony about the methods and procedures actually used in selecting whom to interview and in conducting the interviews is pertinent and required. (*Squirt Co. v. Seven-Up et. al.* 207 USPQ 9, 27. *American Home Products v. Johnson & Johnson et. al.* 196 USPQ 489.) Each person testifies regarding what he or she did, or knows from direct observation, con-

cerning how a survey was conducted and analyzed. Direct and cross-examination deal with the expert standing of the responsible individuals, the experience of the survey organization, the care with which the survey was administered, and the soundness of the analytical methods that were used. In this way, the integrity of the adversarial system is protected, without having to cross-examine survey respondents.

The use of survey data as legal evidence makes confrontations inevitable between research professionals who provide expert testimony to contending parties. The cross-examination of experts can be stressful, especially when professionals who know and respect each other find themselves testifying for opposing parties. I recently was in such a position, testifying about an employee survey I had worked on for Sears, Roebuck. That survey had been introduced as evidence in a suit instituted against Sears by EEOC. A methodological issue arose as to which of two questions correlates better with behavior and, therefore, should be used as an indicator of how interested non-commissioned salespeople are in being promoted to commission sales. No external criterion for validation was available, so the court was faced with the problem of deciding which of two experts, each of whom respected the other's credentials, was correct. Seeking an empirical resolution to this confrontation, I eventually suggested that by cross-tabulating the two items, a quasi-scale could be formed to differentiate those with a real, active interest in promotion to commission sales from those with a passive quasi-interest.

In accord with standard non-jury trial procedures, the correctness of my suggestion is to be decided not by further research but by the judge. This is illustrative of the principle that confrontations between survey researchers who serve as experts to contending parties are decided not by their peers, which is something that scientists should expect and welcome, but by non-scientists who use their own criteria. If the adversary system for determining truth were to be abandoned, other methods for resolving differences between expert witnesses might be adopted. But, the likelihood of that happening seems small indeed.

In any event, overcoming the exclusion of survey data as legal evidence has not been achieved through a dispensation from the hearsay rule and the strictures of the adversarial procedure for reaching "truth." Courts are still concerned that "without cross-examination there is no testing of sincerity, narrative ability, perception, and memory (nor) showing whether they [respondents] were influenced by leading questions, the environment in which the questions were asked, or the personality of the investigator." (*American Luggage Works v. United States Trunk Co.*, 158 F. Supp. 50, 53). The courts have not decided to change the institutional norms that govern our legal system but, rather, how those norms should be applied to survey data. Recognizing that interview surveys have created a new kind of empirical information about the world we live in, courts have devised ways for applying traditional legal criteria for judging that information's worth as evidence.

Surveys and trademark litigation

The hearsay rule has not been the only barrier to the admission of survey data as legal evidence.

Although individual judges do have discretion as to whether they will admit survey data, and how much weight they will give to that data, they are governed by legal principles that encompass, but go far beyond, methodological criteria of concern to survey researchers. In fact, the process whereby survey data are admitted or excluded is governed by rules that may have little, if anything, to do with what researchers consider most important for evaluating a survey's scientific worth. Court rulings in trademark litigation, possibly the single most common type of litigation in which survey data are used, illustrate how important it is to consider legal, and not only methodological, principles when designing and analyzing a survey for litigation.

Trademark disputes typically relate in some manner to the question of "whether there is any likelihood that an appreciable number of ordinarily prudent purchasers are likely to be misled, or indeed simply confused, as to the source of the goods in question." (*Mushroom Makers, Inc. v. R.G. Barry Corp.*, 580 F. 2d 44, 47 199 USPQ 65, 66-67). What has come to be called the "Polaroid formula" established eight criteria for assessing the likelihood of confusion: The strength of the mark, degree of similarity between the two marks, proximity of the products, the likelihood that the prior owner will bridge the gap, actual confusion, the defendant's good faith in adopting its own mark, the quality of the defendant's product, and the sophistication of buyers. (*Polaroid Corp. v. Polarad Electronics Corp.*, 287 F. 2d 492, 495, 128 USPQ 411, 412-13, 2d Circ.)

Courts have repeatedly ruled that the proper application of this formula requires determining (1) whether the mark has a primary or arbitrary meaning, (2) in the absence of an arbitrary meaning whether it has an acquired or secondary meaning, (3) the closeness of products in the selling situation and whether they utilize the same trade channels, (4) whether the products are sold to the same class of customers, and (5) the care exercised by customers when buying products of the class involved. (*Philip Morris Inc. v. R.J. Reynolds Tobacco Co.* 188 USPQ 289, 292-3; *RJR Foods, Inc. v. White Rock Corp. et. al.* 201 USPQ 579, 581-582; *Ives Laboratories, Inc. v. Darby Drug Co., Inc. et. al.* 206 USPQ 238, 242-243; *Monsieur Henri Wines, Ltd., et. al. v. Duran* 204 USPQ 604-605). The relevance of survey data to each of these criteria differs considerably.

In trademark litigation, unless a survey is specifically designed to meet the above criteria, the likelihood is strong that it will be excluded from evidence -- no matter how well designed it may be. Even if admitted, limited weight, if any, will be given to a survey if more direct evidence on the points at issue is available. For example, non-survey evidence regarding price, ingredients and manufacturing methods, and usual trade outlets -- characteristics used to ascertain product quality -- are ordinarily readily available. Consequently, attempts to establish the quality of a product through survey data are seldom successful. Only if other sources of information are not available, or if a survey can cast doubt on those other sources, is a survey likely to be of value in establishing a product's quality.

Similarly, the strength of a mark can often be demonstrated by the length of time it has been

used, the consistency of its use, and the size of advertising budgets, making survey data superfluous. (*Squirt Co. v. Seven-Up Co.*, 207 USPQ, 9-10). Survey researchers who do not appreciate such legal considerations when designing a survey will be frustrated by the seemingly arbitrary disregard of surveys that a marketing manager might find well-designed for his purposes.

Courts may determine customer sophistication on the basis of such non-survey data as price, how specialized a product is, and whether such factors as occupation give buyers an expert status (*American Luggage Works, Inc. et. al. v. United States Trunk Co., Inc.* F. Supp. 50, 53). Nonetheless, when surveys provide relevant information that is not available from other sources, courts have been receptive to their use for assessing customer sophistication. In these instances, survey data have fared best when they describe behavior, for example, the frequency with which a product is purchased, brand loyalty, knowledge of specific product qualities, specific considerations that determined buying decisions, and how much comparative shopping is done for a product. Such data cast light on customer sophistication by focusing on the "live market situation," which courts consider preferable to asking respondents for statements of opinion.

Although measures of behavior are of more interest to courts than measures of opinion, the meaning of the behavior must be apparent to the court, and not imputed. A court's distinction between "apparent" and "imputed" can be far more rigorous than a marketing manager's. For example, one judge decided that the fact that a customer bought one brand but named another with a similar name when asked what brands had been bought does not necessarily demonstrate confusion. To demonstrate that there was confusion, he ruled, evidence is needed that explicitly identifies why that brand was bought. (*Squirt Co. v. Seven-Up Co.*, 207 USPQ, 14).

Despite the preference for behavioral rather than opinion data, courts have been more receptive to surveys that describe psychological states than to those that describe objective reality. "What interviewees said is offered to show not the truth of what the interviewees said but to show their state of mind." (*American Luggage Works v. United States Trunk Co.*, 158 Supp. 50, 53). This is why the most common use of surveys in trademark litigation is when the meaning of a mark or confusion between marks have been at issue. In these instances, it is the consumer's understanding of symbols that is at issue, and courts have accepted the principle that such understanding can best be ascertained by talking to consumers. For example, they have ruled that interviews can provide direct evidence as to whether consumers attribute a primary or secondary meaning to a mark.

The direct relevance of survey data to litigation about a mark's meaning and likelihood of confusion can be a trap for the unwary user. Surveys which do no more than demonstrate the possibility that some people may be confused or impute a certain meaning to a mark will probably be rejected by a court. What is required is that respondents represent actual or demonstrably prospective buyers of the specific product class, and that the survey test respondents as closely as possible to real

buying situations:

If the interviewee is not in a buying mood but is just in a friendly mood answering a pollster, his degree of attention is quite different from what it would be had he his wallet in hand. Many men do not take the same trouble to avoid confusion when they are responding to sociological investigators as when they spend their cash. (*American Luggage Works v. United States Trunk Co., Inc.* 158 F Supp. 50, 116).

Satisfying legal logic

Marketing managers are as concerned as courts that survey data relate to real life behavior. But when under pressure to make a decision without all the information he would ideally want, a marketing manager will often consider possible implications of survey data that a court would rule inadmissible. It is incumbent upon the survey researcher to keep this difference in mind when developing research designs for surveys that will be used in litigation. The researcher must develop designs that are meticulous in their attention to the critical logic to which the survey's results will be subjected. If this is not done, a survey that a marketing manager might well consider useful might be rejected by a court as speculative.

A survey I directed for the National Commission on Egg Nutrition illustrates this point. The survey's objective was to measure how many people avoid eating eggs because they believe the high cholesterol content of eggs increases the risk of developing heart disease. The survey results were to be entered as evidence in action that the Federal Trade Commission had taken against the NCEN with respect to an advertising campaign. That campaign claimed that at the time (1975) the alleged health risk in eating eggs had not been scientifically established. The NCEN claimed a First Amendment right to voice its opinion in advertising. (*Commission on Egg Nutrition V. Federal Trade Commission, U.S. Court of Appeal, 7th Circuit, 570 F. 2, 157*).

I felt that standards of admissibility would not be met if the survey only determined how many people believe the high cholesterol content of eggs made eating them a health hazard and gave that as a reason for not eating eggs. Such results would probably meet the needs of a marketing manager who wanted to know if fear of cholesterol was inhibiting egg consumption. However, our problem was not only to establish that health concerns were probably inhibiting egg consumption. We also had to satisfy a court that we had both (a) measured how many people were not eating eggs specifically because of concern about cholesterol and (b) measured the state of mind of the relevant segment of the public as closely as possible to real life situations.

With these considerations in mind, the questionnaire was structured as a survey of food consumption habits, with health concerns and eggs featured only at the end. The first part of the question-

naire was designed to obtain a detailed behavioral description of consumption of all types of food and not only eggs. This provided an measure of the incidence of egg consumers uninfluenced by reference to motivation or emphasis on eggs. The next section investigated attitudes toward different foods, probing for reasons why some are preferred and others avoided -- but with no special focus on health or eggs. This provided an unaided measure of health considerations as a reason for not eating eggs.

A series of questions was then asked about awareness of and knowledge about cholesterol, saturated fats, the foods that are high in each, and the relation of each substance to heart disease. Questions were then asked specifically about (a) awareness of claims that eggs are a possible factor in heart disease because they are high in cholesterol and (b) the state of scientific evidence about those claims. These last two sets of data measured the knowledge base for those who limit their egg consumption. It was then possible, analytically, to form an estimate of how many people excluded eggs from their diet, or limited their egg consumption, because of health concerns, rather than for other reasons such as life style and taste. It was also possible to relate this estimate to perceptions of the scientific standing of the hypothesis that dietary cholesterol is a hazard to everyone and not only those with a high risk of heart disease. The results of this survey were accepted with virtually no challenge.

An interesting footnote to this 1975 survey is that its results suggested that pork producers might be facing an incipient marketing problem because of public awareness of the high cholesterol and saturated fat content of pork. The data on this point were by no means definitive, and were unlikely to be convincing to a court. However, some of the NCEN people commented that although pork consumption did not appear to have been affected as yet by these perceptions, if they were in the pork business they would be very concerned about the need to counteract possible future impact. This illustrates the contrast between legal and marketing attitudes toward the use of the same survey data.

Conclusion

The admissibility of a survey is always to some degree problematic since the decision is made by individual courts. Methodological soundness is always an important criterion for admissibility, but it is not sufficient. Survey researchers need to appreciate that judicial decisions on the admissibility and weightiness of survey data relate to legal considerations founded on a rational system of thought. This is a point ignored by many survey researchers who complain that decisions are too often based on irrelevant, even capricious, considerations. The successful use of survey data as legal evidence depends on learning to function within the rules of the legal system, and not by trying to circumvent them.