

# Uniqueness of Survey Evidence

Hans Zeisel

Follow this and additional works at: <http://scholarship.law.cornell.edu/clr>

 Part of the [Law Commons](#)

---

### Recommended Citation

Hans Zeisel, *Uniqueness of Survey Evidence*, 45 Cornell L. Rev. 322 (1960)  
Available at: <http://scholarship.law.cornell.edu/clr/vol45/iss2/6>

This Article is brought to you for free and open access by the Journals at Scholarship@Cornell Law: A Digital Repository. It has been accepted for inclusion in Cornell Law Review by an authorized administrator of Scholarship@Cornell Law: A Digital Repository. For more information, please contact [jmp8@cornell.edu](mailto:jmp8@cornell.edu).

# THE UNIQUENESS OF SURVEY EVIDENCE

Hans Zeisel†

*"Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized . . . ."*\*

Barely a decade ago, in a trade-mark confusion case involving girdles for young ladies,<sup>1</sup> Judge Jerome Frank playfully supported his dissenting opinion by a survey which he himself called "not satisfactory." In fact, it was probably the worst survey ever made. He simply solicited his daughter's and her girl friends' opinions.

Nevertheless, Judge Frank had a point. Sadly remarking that "neither the trial judge nor any member of this court is (or resembles) a teen-age girl," he felt that where the opinions of teen-agers are at issue nothing but "information directly obtained" from them would suffice. Regretting that he had no "staff of investigators like those supplied to administrative agencies" he proceeded on the sound principle that any relevant information is better than none.

By now, the science of making surveys has come into its own, and with increasing frequency the courts are called upon to decide questions relating to the admissibility and use of survey evidence. Administrative agencies, not bound by the technical rules of evidence, will usually admit a survey and then evaluate its probative power.<sup>2</sup> In court trials, especially jury trials, the issue of a survey's worth will first arise when admissibility is decided on. But whatever the formal occasion, the evidential value of surveys is at issue. This paper is an effort to contribute towards a better understanding of survey evidence, its value and its limitations, within the legal framework which the courts have begun to develop.<sup>3</sup>

---

† See Contributors' Section, Masthead, p. 347, for biographical data.

\* *Frye v. United States*, 293 Fed. 1013, 1014 (D.C. Cir. 1923).

<sup>1</sup> *Triangle Publications v. Rohrlich*, 167 F.2d 969, 976 (2d Cir. 1948).

<sup>2</sup> . . . It has long been settled that the technical rules for the exclusion of evidence applicable in jury trials do not apply to proceedings before federal administrative agencies in the absence of a statutory requirement that such rules are to be observed. *Opp Cotton Mills v. Administrator*, 312 U.S. 126, 155 (1941).

<sup>3</sup> A review of the law and literature through 1953 will be found in Lester E. Waterbury, "Opinion Surveys in Civil Litigation," a paper delivered at the 1952 meeting of the American Association for Public Opinion Research, 17 *Public Opinion Quarterly* 71 (1953); a revised version in 44 *Trade Mark Rep.* 343 (1954); Note, 66 *Harv. L. Rev.* 498 (1953). A more recent and more detailed, but less analytical, inventory can be found in Barksdale, *The Use of Survey Research Findings as Legal Evidence* (1957). A review of the potentialities and shortcomings of public opinion research, important for both the lawyer and the survey practitioner, will be found in Blum and Kalven, "The Art of Opinion Research: A Lawyer's Appraisal of an Emerging Science," 24 *U. Chi. L. Rev.* 1 (1956). Appended (at 64-65)

The unique value of a survey consists in its being the only means of measuring the characteristics of a group where such a measure is at issue, *e.g.*, the proportion of people who confuse two trade-marks, the share of a market held by a competitor, the level of wages paid in an industry, the amount of timber burned in a forest fire, the proportion of items manufactured below specific standards.

The two survey aspects which raise major legal issues are: (1) whether the information is gathered through verbal statements made by the survey interviewees, or only through observation, counting or measuring by the field worker, and (2) whether the survey is based on a census enumeration involving *all* members of the group, or merely on a sample taken from that group.

The first problem raises the issue of hearsay. The second raises the question as to how interviewing a few hundred people, or examining a few specimens, can produce correct conclusions concerning many hundred thousands of people or a whole shipload of merchandise.

Accordingly, for our purpose, surveys may be classified into four groups: (1) *Census* surveys *not* involving *verbal statements*, *e.g.*, surveyors measuring the acreage of an area, and bookkeepers or accountants determining the amount of a designated type of expenditure; (2) *Census* surveys involving *verbal statements*, *e.g.*, the standard job of the U.S. Census; (3) *Sampling* surveys *not* involving *verbal statements*, *e.g.*, the job of the Food and Drug Administration in examining shipments of merchandise, a survey of car license plates of patrons to determine their geographic distribution,<sup>4</sup> and certain phases of accounting work;<sup>5</sup> and (4) *Sampling* surveys involving *verbal statements*, *e.g.*, public opinion polls and similar interviewing operations.

The major legal difficulties are compounded in surveys of type (4). These public opinion poll surveys, therefore, will form the center of this discussion.

The law has not yet developed a general rule for the treatment of all survey evidence, although it is recognized that only two types of surveys are clearly acceptable. Survey evidence produced by the U.S. Census, although it is hearsay and frequently based on sampling rather than "census" operations, is clearly admissible. The other type of survey is admissible, not because of its distinguished authorship, but because of the nature of the collected information. If the interview answers are not

---

are comments by two distinguished social scientists, Paul F. Lazarsfeld and Samuel A. Stouffer.

<sup>4</sup> This serves to establish the competitive position between retail outlets, or motion picture theatres. See note 73 *infra*.

<sup>5</sup> See *infra* p. 329.

gathered "for the truth of the matter asserted therein" the survey will be accommodated under the state-of-mind exception to the hearsay rule.<sup>6</sup>

Surveys which do not fall into either of these two categories can gain admittance only under the broader, less stringent approach (of which the state-of-mind exception is but a special case) which weighs the necessity of utilizing the evidence against the dangers of hearsay. Unless, therefore, a survey exception to the hearsay rule emerges at some future time, the admissibility of survey evidence, as a rule, will hinge upon the assessment of its worth.

In two recent cases, the sharp dividing line between state-of-mind surveys and other surveys was softened. In one, a state-of-mind survey was admitted because there seemed to be sufficient safeguards for admitting that particular hearsay evidence.<sup>7</sup> In the other, a survey was admitted "for the truth of the matter asserted" in spite of the fact that it *was* made, because the hearsay dangers were found to be negligible compared with the value of the promised information.<sup>8</sup>

The privileged status which the law accords to the operations of the United States Census provides an appropriate introduction for our discussion because here, in a curious way, the law has solved, expressly or by implication, all the major perplexities created by surveys and sampling operations. Thereafter, we will consider the problems of sampling and hearsay as they arise with respect to other surveys. Our discussion will suggest that the true difficulties with survey evidence do *not* arise from its being based on sampling operations or on hearsay. Rather, these factors will emerge as the legal justifications for avoiding difficulties that can probably be met more directly.

#### THE UNITED STATES CENSUS

Courts have not hesitated to take judicial notice of Census data, and some state statutes make such data competent "prima facie evidence" as a matter of law.<sup>9</sup> This is so although all Census data are technically hearsay, transferred as they are from the original respondent through a long chain of staff workers. Furthermore, much Census data are based on sample operations rather than a complete enumeration of all units.

One explanation for the fact that the sampling issue has never been raised in connection with Census data is, perhaps, that so few people know that many Census data are based only on samples. For instance,

---

<sup>6</sup> See note 37 *infra*.

<sup>7</sup> See note 39 *infra*.

<sup>8</sup> See note 40 *infra*.

<sup>9</sup> 5 Wigmore, *Evidence* § 1671 n. 11 (3rd ed. 1940).

the Census income distribution for the United States population and its subgroups is based on the questioning of every fifth household head. Similarly, such household data as the "proportion of homes having mechanical refrigeration" are based on replies from inhabitants in every fifth dwelling unit. Yet the Census Bureau considers these findings not significantly less reliable than actual "census" data. It is an obvious waste of resources to ask a question of all respondents, if questioning every fifth will yield results of sufficient, if not greater, accuracy.<sup>10</sup>

Even if the sampling nature of some Census data were better known, it is improbable that it would be used to impeach their accuracy. This may be inferred from the emphatic court endorsement of the counting procedures of an official institution of much lesser standing. With respect to a sampling examination conducted by the Federal State Seed Laboratory in Alabama, a court said:

Reports which are of a public nature and taken under competent authority to ascertain a matter of public interest are admissible in evidence against all the world.<sup>11</sup>

The survey operations of the Census have still another privilege vital to the success of any continuous interviewing operation. By statute, the information given to the Census enumerator is treated as privileged, which makes it impossible to verify the obtained information by calling interviewees as witnesses in court.

For survey operations not conducted by the U.S. Census or similar public bodies, all three objections—that they are based on samples, that they constitute hearsay, and that there is a reluctance to abandon the anonymity of their interviewees—provide serious obstacles for the law. One must ask, therefore, why a status given to the U.S. Census is being withheld from other survey operations? Whatever the legal formula, the true answer must lie in the confidence which the Census survey peculiarly commands. This confidence derives from the disinterested character of the Census operation and from the trust in its expertness.<sup>12</sup> It is sustained by the great simplicity and the factual character of most Census questions, which leave little room for error.<sup>13</sup>

---

<sup>10</sup> See note 25 *infra*.

<sup>11</sup> *E. K. Hardison Seed Co. v. Jones*, 149 F.2d 252, 257 (6th Cir. 1945).

<sup>12</sup> Compare, however, this passage in Judge Caffey's opinion in the *Alcoa* case:

. . . I conceive of no reason for discriminating between private individuals and Government employees in formulating . . . what constitutes inadmissible hearsay.

*United States v. Aluminum Co. of America*, 35 F. Supp. 820 at 825 (S.D.N.Y. 1940).

<sup>13</sup> Still another sample survey which is admitted almost daily without much legal argument in our courts, and probably without awareness that it is a sample survey, is the mortality table, submitted as proof of the average life expectancy at any given age, subject to modification by special trials of the particular individual.

## SAMPLING PRINCIPLES

Most surveys are based on samples because, practically speaking, there is no other way of conducting them. There are even a few situations where a complete census enumeration is not merely impracticable but impossible because the census would destroy the entire group to be measured. This is inevitable whenever the examination involves the destruction of the examined unit, as in most laboratory examinations of food and drugs.<sup>14</sup> Similarly, sampling naturally recommends itself whenever the universe (the technical term for the group from which a sample is taken) is either physically unavailable or remains undetermined as to its ultimate size. Estimating the value of burned timber is an example of the first;<sup>15</sup> estimating the potential loss of bus fare from a projected change-over to one-way traffic is illustrative of the second.<sup>16</sup>

The most frequent and, in the long run, most important advantage of samples is their overwhelming economy as compared to the often prohibitive costs of the census operation. While most universes would permit a census survey in theory, its costs, in terms of money and time, as a rule, are prohibitive. In an early attempt to approach something like a census survey in a Federal Trade Commission case, one of the commissioners voiced this eloquent protest:

I want to register my protest at the way in which this case was conducted. About a thousand witnesses . . . were permitted to testify as to whether the use of the word 'Castile' when applied to a soap not made exclusively of olive oil, had the tendency to deceive the public . . . [T]his piling up of cumulative evidence is an inexcusable outrage on the public . . . The attorneys and the trial examiner traveled throughout the country for the purpose of taking the testimony of such witnesses. About 700 such were subpoenaed to testify at Spokane . . . [T]estimony has caused the F.T.C. to waste hundreds of thousands of dollars.<sup>17</sup>

Judge Wyzanski, in the *Shoe Machinery* Case, notes that "If anti-trust trials are to be kept manageable, samples must be used."<sup>18</sup>

---

<sup>14</sup> . . . [a] test(s) . . . results in the article's being rendered useless . . . [I]nspection and condemnation on the basis of samples tested is clearly contemplated by the Act [21 U.S.C.A. 334(c)].  
United States v. 43 and one-half Gross Rubber Prophylactics, 65 F. Supp. 534, 536 (D. Minn. 1946).

<sup>15</sup> [The witness] took four sample plots in each 40 acre subdivision and estimated the destroyed merchantable timber . . . We are not convinced that any improper elements entered into the estimates of the witness. . . .  
Feather River Lumber Co. v. United States, 30 F.2d 642, 643 (9th Cir. 1929).

<sup>16</sup> The Court accepted a sample survey and the opinion evidence of an expert witness on the ground that it was "dictated by necessity, there being no other way to establish the probable loss of fares." Eighth Avenue Coach Corp. v. City of New York, 170 Misc. 243, 251, 10 N.Y.S.2d 170, 181 (Sup. Ct. N.Y. County 1939), aff'd, 286 N.Y. 84, 35 N.E.2d 907 (1941).

<sup>17</sup> From Commissioner Humphrey's dissent in Matter of James S. Kirk & Co., 12 F.T.C. 272, 289 (1928).

<sup>18</sup> United States v. United Shoe Machinery Corp., 110 F. Supp. 295, 305 (D. Mass.

Occasionally, a situation may arise where the universe to be examined is finite and small enough so that the law, by insisting upon complete enumeration, does not impose an impossible burden on the litigant. In a California case, involving suit for overpayment of a local sales tax, plaintiff offered an estimate of the overpayment based upon a sample, but the court insisted upon an introduction of the complete count.<sup>19</sup>

The courts generally are familiar with the fact that a sample must be representative of its universe. To state precisely what is meant by a representative sample, we might start by recalling a most famous instance of what was *not* a representative sample—the more than two-million ballots of the 1936 Literary Digest poll which forecast Landon's victory over Roosevelt. By relying on the voluntary response from people listed in telephone directories, owners of automobile licenses, and its own subscribers, the Literary Digest sample *under*-represented those social strata which were to sweep Roosevelt into his second term. A representative sample, then, is one in which all members of the group have an equal chance of being selected. Such a sample will represent all subgroups more or less proportionately.<sup>20</sup>

The only way to insure an adequate representation is by some form of lottery. A lottery, with *all* members of the universe susceptible of selection, is at the core of every sampling operation. The so-called systematic sample, whereby every *n*th item or name from a list is selected, is but a hidden form of lottery. It is implied that these *n*th units are in fact random selections.<sup>21</sup> The many variants and problems in sample design do not so much derive from the basic and easily solved problem of se-

---

1953), *aff'd*, 347 U.S. 521 (1954). Compare also the strong endorsement of survey evidence in "streamlining the Big Case" Report of Special Comm. of the Section of Anti-Trust Law, Am. Bar. Assoc., Sept. 15, 1958, p. 31.

<sup>19</sup> *Sears Roebuck & Co. v. The City of Inglewood*, reported in Sprowls, "The Admissibility of Sample Data into a Court of Law," 4 U.C.L.A.L. Rev. 222-32 (1957). See also note 34 *infra*.

<sup>20</sup> Sometimes, however, if one wants to measure characteristics of a relatively small sub-group it will be necessary to over-represent that sub-group in the sample (over-sampling) so as to have a sufficiently large cross-section of the sub-group. If one wanted to compare, for instance, households with and without automobiles, one might sample every non-owner but only every fourth owner, so as to have equally sized samples of both groups.

<sup>21</sup> Judge Wyzanski used a sample of similar design:

At the Court's suggestion, the Government took and offered in addition to the OMR's [Outside Machine Reports] depositions [a sample] of 45 shoe manufacturers operating 55 factories. The Court arbitrarily selected from a standard directory of shoe manufacturers, the first 15 names that began with the first letter of the alphabet, the first 15 names that began with the eleventh letter of the alphabet, all 8 of the names that began with the twenty-first letter of the alphabet, and the first seven of the names that began with the twenty-second letter of the alphabet. This sample covers 3 per cent of the shoe manufacturers. The sample includes small and large factories, and concerns manufacturing shoes according to substantially the most popular shoe manufacturing processes.

*United States v. United Shoe Machinery Corp.*, 110 F. Supp. 295, 305. (D. Mass. 1953).

Such clustering, however, is not without danger since it may coincide with natural clusters on the list, such as the names beginning with "Mc" or with "Rosen."

lecting a representative sample,<sup>22</sup> but from the secondary difficulty of selecting a sample with *maximum efficiency*, that is, with minimum costs.

#### LIMITATIONS UPON THE ACCURACY OF SAMPLE MEASUREMENTS

The price to be paid for accepting a sample instead of a census enumeration is a certain degree of uncertainty. This might seem a shocking admission in an argument for the acceptance of sampling evidence. But this shortcoming, as we shall see, is mitigated, first, by the fact that the degree of uncertainty can be reduced to any desired magnitude (though never to zero); second, because the law is rarely in a position to insist on certainty;<sup>23</sup> and third, and most important, by the fact that census operations, too, are open to error. Experts insist that a sample survey, in spite of its inherent uncertainty, can, at times, be more accurate than the corresponding census operation.<sup>24</sup>

Let us first clarify the nature of this uncertainty in sample measurements. Suppose we found that 68 per cent of the housewives in a sample had bought at the XX food store chain. What can we say about the corresponding percentage in the universe from which this sample was drawn? Ignoring the technical niceties, the relationship is expressed as follows: A range is set, extending in both directions around the sample measurement, e.g., from 62 to 74 per cent, also written as  $68 \pm 6$  per cent. The probability that the true (universe) measure, if determined by a census, will fall within this range is then determined. If in this particular instance the sample measurement of 68 per cent was derived from a sample of 5,400 housewives, the odds will be 99 to 1 that the

<sup>22</sup> The Federal Seed Act Regulations, 7 C.F.R. §§ 201.208-19, and especially §§ 201.210-13 (1959) contain explicit sampling instructions and thereby acknowledge this underlying principle.

<sup>23</sup> . . . [A]lthough, admittedly such summaries are not, and could not be, absolutely complete, they are the most accurate and most dependable data in existence to establish this fact.

Hughes Tool Co. v. Ford, 114 F. Supp. 525, 542 (E.D. Okla. 1953).

<sup>24</sup> The paradox is easily explained. Census operations are, by definition, gigantic tasks, necessitating the hiring of a great staff, the quality of which eludes control. It is suggested that population sampling surveys, in certain respects, would produce more accurate results than the population 'census.' Some of the Census inaccuracies are well known: Reported age distributions have a slight tendency to bulge at the round (10 year) intervals; transient (hotel, etc.) populations are underrated; babies under 1 year similarly are underrated—at the expense of babies who are reported as having reached their first year.

For the array of the survey problems the Bureau of the Census is confronted with, compare Hauser, "Labor Force and Gainful Workers—Concept, Measurement, and Comparability," 54 J. Sociology 338 (1949).

Compare, also, the following reference from a case before the Illinois Commerce Commission:

The Company's evidence with respect to depreciation . . . was based primarily upon a sample of the Company's property . . . . It was pointed out that the use of a sample tends toward greater accuracy in the final conclusions, inasmuch as a reasonably intensive inspection of the entire plant is virtually impossible considering the limitation of time and money that should reasonably be imposed.

Illinois Bell Telephone Co., 7 P.U.R.3d 41, 606 Ill. Commerce Comm'n 493, 506 (1955).

statement is correct that the true universe percentage will fall within the indicated limit. Given a certain sample, the odds will be lower where the narrower limits of the range are posted. In our example, the odds will be about 19 to 1 that one is correct in asserting that the true value will fall within  $68 \pm 4$  per cent, and only 2 to 1 that they will fall within  $68 \pm 2$  per cent, *i.e.*, between 66 and 70 per cent.

A special case of this estimating procedure is acceptance sampling, a customary practice in quality control of production and accounting. The task is to determine, by means of a sample and in terms of agreed-upon odds, whether the proportion of faulty items in the universe does or does not exceed a specified tolerable minimum. If no faulty item is tolerable, sampling must, of course, never be resorted to, as that item might be hidden in the non-sampled part of the universe.<sup>25</sup> But if some small proportion of faulty items is tolerable, however small, a sample can be designed to assure confidence odds for such a margin.<sup>26</sup>

This situation requires the survey planner to determine the necessary accuracy of the sample measurement and to design his sample accordingly. The accuracy with which a sample result can be projected depends, as a rule, on two factors: the *absolute size* of the sample and the *variance* of the property to be measured. The widely held belief that the accuracy of a sample is connected with its relative size to the universe is mistaken.<sup>27</sup> A sample smaller than 1 per cent, taken from one universe, can be much more reliable than one comprising 10 per cent of another. To determine with equal accuracy the average age of the population of New York City and of Peoria, Illinois will require samples of equal size.

But if the *size* of the universe is immaterial, its *variance* is of great importance. Assume we wanted to determine the mean age of two groups of people and wanted to be sure of certain minimum odds that the sample

---

<sup>25</sup> To think otherwise would be like the person who, in counting a bundle said to contain 1000 one dollar notes, stops satisfied at 996 and exclaims: "If it was all right up to now, the rest will be all right too."

<sup>26</sup> Sampling, as a legitimate part of the accountant's practice, is acknowledged in *Ultramares Corp. v. Touche*, 255 N.Y. 170, 191-92, 174 N.E. 441, 449 (1931).

The defendants [charged with having falsely certified a balance sheet] attempt to excuse the omission of an inspection of the invoices proved to be fictitious by invoking a practice known as that of testing and sampling. A random choice of accounts is made from the total number on the books, and these, if found to be regular . . . are taken as a fair indication of the quality of the mass. . . . Verification by . . . sample was very likely a sufficient audit as to accounts regularly entered upon the books in the usual course of business.

The court goes on to say that the sample in the present case did not excuse the omission, not because sampling in itself is improper, but only because the crucial entry in the ledger of assets, plainly interpolated and suspicious on its face, was excluded from the sample. For technical details on sampling in accounting work, see Vance and Netter, *Statistical Sampling for Auditors and Accountants* (1956).

<sup>27</sup> Only in those rare cases where the sample becomes a sizable proportion of the universe, say 20 per cent or more, do adjustments for approaching the size of the universe need to be made in the formula.

result be not more than 3 years off the true mean age of the sampled group. Suppose, further, that one group to be sampled is the *general* population of a city and the second is the residents in all the *old age homes* of that city. Since age in the general population varies from zero to approximately one-hundred, the general population will, for the same degree of accuracy, require a larger sample than the old age home population, whose age varies only between seventy and one-hundred.

The inverse relationship between variance and sample size is most visible in the sampling of well mixed liquids; a drop of blood will provide an accurate count of the blood cells per volume, because this percentage varies little from drop to drop. There is, indeed, no more homogenous universe than a well mixed liquid. Hence, even a very small sample might yield an accurate measurement. Yet the following generalization is not warranted:

[One of the] prerequisites necessary to the admission in evidence of samples [is] that the mass should be substantially uniform with reference to the quality in question. . . .<sup>28</sup>

The court errs here, for also a universe with low homogeneity permits sample measurements of great accuracy; it merely requires a correspondingly larger sample. Uniformity or homogeneity or variance—the three terms are used here interchangeably—is a matter of degree and can be compensated by appropriate sample design and size.

Two points emerge; first, that the degree of accuracy of any sample can be increased to the desired limit; and second, that the degree of accuracy with which the sample is related to its corresponding value in the universe can be stated with precision. Expert knowledge of the sampling error is as unambiguous as the knowledge of the relationship between a circle's radius and its circumference.

#### THE SAMPLING ERROR AND THE LAW

Sample measurements raise interesting legal issues. It might seem disconcerting that no sample measurement can be stated with complete confidence in its accuracy. But, as we pointed out, the law is accustomed to dealing with less than perfect evidence. Both standards of proof, "reasonable doubt" and "preponderance of evidence," allow for imperfections. Rather, it is the positive aspect of sample measurements that raises new issues, *i.e.*, the possibility of actually measuring the degree of uncertainty or certainty through tolerance limits for the measurement at issue.

One way in which the law might deal with the problem is to accept the

---

<sup>28</sup> E. K. Hardison Seed Co. v. Jones, *supra* note 11 at 256.

range of the sampling error because *any* value within its range would fulfill the immediate legal requirements:

For the evidence is all in one direction . . . . Different methods of weighting the various factors would produce different results; but no reasonable, qualified person would, by any rational process, reach a figure outside that range; and probably most methods would reach a figure close to the middle of it. Moreover, even though this high figure is not attained in every part of the market, nonetheless the figure may be fairly used since United supplies in every significant generic class of shoe machinery, except machinery used in the cement process, and in rubber shoe manufacturing, and of course, excepting dry thread sewing machinery, far more than 50% of the demand. In short, it is not inaccurate in this market to say United has a 75-95% share; and it probably would be accurate to say an approximately 85% share.<sup>29</sup>

Or the result of the measurement, however inaccurate, may clearly fall short of the legal requirement:

. . . [P]laintiff did produce witnesses who testified as to confusion . . . . However, considering that annually more than 70,000,000, rolls of plaintiff's . . . mints are sold, to say nothing of many millions of packages of other flavors . . . it would be extraordinary if some confusion could not be found irrespective of the details of the dress of the package . . . . A new competitor is not held to the obligations of an insurer against all possible confusion.<sup>30</sup>

In these cases the court found the magnitude of the sample measurement so clearly above or below the relevant legal limits that it deemed it unnecessary to put numerical values on these limits. In the following opinion, however, full use was made of the possibility of measuring the accuracy of a sample:

The Commission retained [an expert] . . . to analyze the Company's sample of its plant to determine . . . whether it was fairly representative of the plant as a whole. He testified that in his opinion the sample was fairly drawn and that the chance of it varying as much as 1% from the plant as a whole was negligible. We conclude that the Company's sample should be accepted as a fair cross-section of the plant as a whole.<sup>31</sup>

The courts could go further and specify standards for sample measurements since the sampling error can be reduced at will by enlarging the sample. To be sure, there is a serious cost element to be noted. Since the size of the sample error is inversely related to the *square* of the sample size, the size of the sample must be quadrupled to reduce a sampling error to one-half of its original size.

An interesting solution was proposed in the California case in which

---

<sup>29</sup> Judge Wyzanski in *United States v. United Shoe Machinery Corp.*, supra note 21 at 307.

<sup>30</sup> *Life Savers Corp. v. Curtiss Candy Co.*, 182 F.2d 4, 8 (7th Cir. 1950).

<sup>31</sup> *Illinois Bell Telephone Co.*, supra note 24.

suit was instituted for overpayment of a local sales tax. Plaintiff offered an estimate of the overpayment based on a sample survey as follows: \$28,250 with 2 to 1 odds that the true (census) value would fall between \$23,150 and \$30,350, and 19 to 1 odds that it would fall between \$23,950 and \$32,450. Plaintiff voluntarily reduced his claim to \$27,000, thereby suggesting his willingness to absorb the greater part of the sampling error. The court insisted on a complete census count, however, only to discover that it yielded \$26,750, a sum within 1 per cent of the claimed amount.<sup>32</sup>

In most situations, however, the universe to be measured will be so large that the court will not have the choice between a sample measurement and a census, but rather, would either have to accept a sample measurement or have no measurement at all. In the latter situation the relatively inaccurate sample measurement may provide better evidence than the law now possesses. That a sample value may be inaccurate by some small margin of error could be of less importance than the fact that it *can* provide a measurement as accurate as it does. This possibility becomes particularly important when the measurement refers to the core of the litigated issue, and hence becomes a measure of the soundness of the court's judgment or verdict. Two examples may be cited.

The owner of a gambling establishment, at his trial<sup>33</sup> for income tax evasion, contended that his lottery wheel had retained, over the years in question, some 11 per cent of the placed bets. The prosecution introduced expert evidence to the effect that the laws of probability suggest that the wheel (with its specific game rules) must earn, in the long run, some 22 per cent of the gross intake, that is, about twice the amount for which the defendant argued. The prosecution might have gone one interesting step further and offered evidence in the following form: "The odds that the defendant's statement is true—that the wheel earned not more than 11 per cent—are one in a thousand" (or whatever the odds were, computed from the rules of the game and the actual number and size of the placed bets).

The use of blood tests in paternity suits provides another example. It appears that, because of the possibility of accidental mutations, their evidentiary value is not completely foolproof. On the average, it is estimated that the blood test may err, by excluding the actual father from parenthood, in one birth in 10,000. Because the paternity cases which

---

<sup>32</sup> *Sears Roebuck & Co. v. The City of Inglewood*, supra note 19. The issue posed in that case could provide a more general rationale for deciding the degree of accuracy that should be required of a sample. The answer could come from asking these two questions:

(1) How much would it cost to increase the accuracy of the sample to a specified limit?  
(2) What would be the dollar equivalent of this gain in accuracy for the issue at hand?

<sup>33</sup> *United States v. Sanders Scott*, 55, 118 (7th Cir.) (unpublished).

reach the courts are more likely to be among the few where the test fails, it is estimated that the test would lead the court into error about 1 time in a hundred. The question arises, therefore, whether, in view of this fallibility, the blood test can still be accepted as incontrovertible proof. The Danish jurist who drew attention to this particular aspect of the problem suggests that blood tests be accepted as such proof. To permit their impeachment by witness testimony could only increase a ratio of error that is otherwise kept to 1 in 100.<sup>34</sup>

#### THE HEARSAY BARRIER

Since opinion surveys reflect statements made by third persons to an interviewer who, in turn, has related them to the analyst who may appear as an expert witness in court, such evidence is clearly hearsay. But the courts have developed a line of authority which exempts surveys from the hearsay rule if the declarations are reported not "for the truth of the matter asserted therein," but as expressions of the interviewee's state-of-mind. Confusion of names and trade-marks provide the leading examples:

The hearsay objection is unfounded . . . . [T]he statements of the persons interviewed were not offered for the truthfulness of their assertions . . . [T]hey were offered solely to show as a fact the reaction of . . . the public. . . . Only the credibility of those who took the statements was involved, and they were before the court.<sup>35</sup>

It is true that some of the hearsay dangers are reduced by an approach of this kind; but they are by no means absent. Thus, while one may be inclined to welcome any opening for the acceptance of survey evidence, one cannot wholeheartedly endorse this particular one.

The state-of-mind doctrine has two disadvantages. It may make the courts overlook technical pitfalls which are more likely to occur in state-of-mind surveys than in other types of surveys. For example, the degree of confusion of two trade-marks is, as a rule, the sum of two forms of confusion: the specific confusion of the two trade-marks in question plus the general confusion that will obtain for *any* comparison in an average group of people under most test conditions. Clearly, unless the degree of specific confusion is great, allowance must be made for the general.

---

<sup>34</sup> Ross, "The Value of Blood Tests as Evidence in Paternity Cases," 71 Harv. L. Rev. 466, 483 (1958). He adds, significantly, that the courts in Denmark are inclined to accept the blood-test exclusion as absolute proof with respect to children born out of wedlock, but show greater reluctance in cases where the paternity of legitimate children is at issue.

<sup>35</sup> United States v. 88 Cases, 187 F.2d 967, 974 (3rd Cir. 1951), cert. denied, 342 U.S. 861 (1951). See also *People v. Franklin Nat'l Bank*, 200 Misc. 557, 105 N.Y.S.2d 81 (Sup. Ct. Nassau County 1951), rev'd, 305 N.Y. 453, 113 N.E.2d 796 (1953), rev'd 347 U.S. 373 (1954); *Household Finance Corp. v. Federal Finance Corp.*, 105 F. Supp. 164 (D. Ariz. 1952). While public opinion polls become almost standard procedure in trade-mark confusion cases, curiously enough, they have never been used to prove (or disprove) the assertion that a trade name had become a generic term.

The more serious difficulty with the state-of-mind rule, however, is that it provides no basis for admitting surveys of *facts*. In the latter type survey, the hearsay danger is less than in surveys of states-of-mind. Surveys which ask the respondents whether they own a gas stove, or carry a certain brand of merchandise in their store, or have reached a certain age, are not admissible under this hearsay exception. Surveys asking for complicated psychological reactions *are* covered. Judge Wyzanski has noted this problem and has decided against the mechanical application of the state-of-mind rule. After reviewing the positions and citing authorities for each,<sup>36</sup> he concludes:

. . . So long as the interviewees are not cross-examined, there is no testing of their sincerity, narrative ability, perception, and memory. There is no showing whether they were influenced by leading questions, the environment in which questions were asked, or the personality of the investigator. But where a court is persuaded that in a particular case all these risks have been minimized, that the answers given by the interviewees are, on the whole, likely to be reliable indicia of their states of mind, that the absence of cross-examination is not prejudicial, and that other ways of getting evidence on the same point are either impractical or burdensome, the testimony should be admitted. . . . In this case these conditions have been met. Accordingly, the hearsay objection is overruled and the testimony of the results of the poll is admitted. . . .<sup>37</sup>

While this position considerably narrows the state-of-mind exception applicable to survey evidence, it must, in the long run, prove sound.

Another case which provides an opening wedge for surveys which cannot be accommodated under the state-of-mind exception is *United States v. Aluminum Company of America*.<sup>38</sup> There, an expert testified to an analysis of the records of 605 test holes drilled to determine the presence and quality of bauxite deposits, although the persons who did the drilling and prepared the reports were not present as witnesses:

Opinion testimony by an acceptable expert resting wholly or partly on information, oral or documentary, recited by him as gathered from others, which is trustworthy and which is practically unobtainable by other means, is competent even though the first hand sources from which the information came be not produced in court.<sup>39</sup>

---

<sup>36</sup> Some authorities have, therefore, concluded that the testimony is not hearsay. *United States v. 88 Cases*, 187 F.2d 967 (3d Cir. 1951); 6 Wigmore, *Evidence* (3d ed.) § 1776; Note 66 *Harv. L. Rev.* 498, 501, 503, note 34. Others, however, have noted that the proffered evidence has some of the dangers of hearsay. See Note 66 *Harv. L. Rev.* 498, 501-502; Morgan, *Hearsay Dangers and the Application of the Hearsay Concept*, 62 *Harv. L. Rev.* 177, 185, 202-203, 206; McCormick, *The Borderland of Hearsay*, 39 *Yale L.J.* 489, 491.

*American Luggage Works v. United States Trunk Co.*, 158 F. Supp. 50, 53 (D. Mass. 1957).

<sup>37</sup> 158 F. Supp. at 53.

<sup>38</sup> 35 F. Supp. 820 (S.D.N.Y. 1940).

<sup>39</sup> *Id.* at 823.

The court cites as precedent the opinion of Judge Learned Hand in *G. & C. Merriam Co. v. Syndicate Publishing Company*:<sup>40</sup>

With respect to the matter, in what impresses me as unambiguous authoritative judicial language, it has been said that 'the requisites of an exception of the hearsay rule' are 'necessity and circumstantial guaranty of trustworthiness.'

The court goes on to clarify the term necessity:

In effect, . . . necessity . . . is not to be interpreted as uniformly demanding a showing of total inaccessibility of firsthand evidence . . . but that necessity exists where otherwise great practical inconvenience would be experienced in making the desired proof. . . .<sup>41</sup>

The opinion, standing alone, is limited. The information in question was not gathered through interviews; hence it involves only one level of hearsay; moreover, it had been gathered in the "ordinary course of business" and only later used for purposes of litigation. Nor does the decision refer directly to surveys but to expert testimony informed *by* a survey. It could, however, be extended to the survey itself.<sup>42</sup>

Let us now consider more closely the dangers of insincerity, faulty narration, perception and memory as they pertain to survey evidence. But let us be sure to see the problem in its precise form: the issue is not whether the reliability of interview response would increase if all interviewees could be examined as witnesses in court, since this is not an available alternative. As a rule, it is not possible to bring the universe or its truly representative sample into court. The customary procedure is to call a number of public witnesses who allegedly are representative of the universe. But a distinguished lawyer with broad experience in this field had this to say about such a procedure: "The poisonous feature of the public witness matter is . . . that all too frequently they are selected not impartially but because they will testify the way the party selecting them wants them to testify."<sup>43</sup> The very fact that such witnesses are arbitrarily selected should render their testimony less credible. Even if these public witnesses would, in fact, give a more reliable response in court than to a survey interviewer, their evidence should be rejected on

---

<sup>40</sup> 207 Fed. 515, 518 (2d Cir. 1913).

<sup>41</sup> *Supra* note 38 at 823-24.

<sup>42</sup> . . . [I]t may very well turn out to be that when the admissibility of survey evidence is finally established beyond doubt, the rationale will be taken from the field of expert testimony.

Waterbury, "Opinion Surveys in Civil Litigation," 44 *Trademark Rep.* 343, 362 (1954).

<sup>43</sup> *Id.* at 347 Waterbury discusses this practice in detail. He notes its appearance in *Stanley Laboratories, Inc. v. FTC*, 138 F.2d 388 (9th Cir. 1943), and *Book-of-the-Month Club, Inc. v. FTC*, 202 F.2d 486 (2d Cir. 1953).

the ground that these individuals do not adequately represent the universe,<sup>44</sup> no matter how many of them are called.<sup>45</sup>

Assuming it were possible to bring a truly representative sample of public witnesses into court, it is doubtful that their answers would always be more reliable than those given to a survey interviewer. The interviewee is in no way connected with the litigants, not even through the tenuous bonds created by being a witness for one side. Moreover, the interviewee will, as a rule, not learn the purpose for which his response is used. In a proper survey routine, to prevent inadvertent disclosure, not even the interviewer is told of the survey's purpose.<sup>46</sup> In addition, since the survey necessarily precedes the trial, less time will have elapsed between the response and the event to be recalled, than between the event and its deposition at the trial. Finally, the court has before it the complete and uniform question schedule in response to which the survey results were obtained. Court witnesses, on the other hand, at times undergo careful individual preparation prior to trial, the form of which does not necessarily come to the court's knowledge.

To be sure, court witnesses may have a heightened awareness of what is at issue and may be more careful and perhaps more perceptive than survey respondents who are completely unaware of the ultimate issues.<sup>47</sup> Cross-examination, too, may prove its value at any time that recollection or narration proves faulty. But even if some of the *individual* survey responses are not, in fact, as totally accurate,<sup>48</sup> the group measurement may still be sufficiently accurate within set tolerance limits.

In summary, therefore, the advantages offered by survey responses should at least suffice to protect such evidence from outright disqualification as hearsay. Moreover, the questions propounded in many cases will

<sup>44</sup> The universe, of course, might be so small and accessible as to permit examination in court of every one of its members.

<sup>45</sup> Hence, it is wrong to believe that the evil of the procedure could be cured by calling a sufficient number of witnesses ". . . except by calling as witnesses so many of the public as to render the task impracticable." *People v. Franklin Nat'l Bank*, supra note 35 at 566. A badly selected sample only becomes worse as it becomes larger. See note 64 infra.

<sup>46</sup> "[The] witnesses were not informed of the purpose of this employment [a shopping\* survey]. . . . Always two went together so that there were two witnesses to each sale." *Oneida, Ltd. v. National Silver Co.*, 25 N.Y.S.2d 271, 286 (Sup. Ct. Madison County 1940).

<sup>47</sup> "Purchases of merchandise are not made in a vacuum with Professor Quiz in charge." *Quaker Oats Co. v. General Mills Inc.*, 134 F.2d 429, 433 (7th Cir. 1943).

<sup>48</sup> Hence, Judge Arnold's concern in *New York Life Ins. Co. v. Taylor*, 147 F.2d 297, 304 (D.C. Cir. 1944) is not quite to the point:

A corporation is engaged in taking a nationwide poll as to the number of members of the Communist Party. In the regular course of that business . . . the interviewer reports that X, Y and Z are Communists, giving excerpts from the conversations to support this opinion. The report would [not] be admissible . . . to make a prima facie case that X, Y and Z are Communists . . . Such evidence might be used in a proceeding for the cancellation of a naturalization certificate.

To be sure, if the unreliability of the individual response passes a certain point, it may invalidate a survey.

be so simple, straight-forward and unambiguous that the hearsay dangers must, in fact, be negligible. Hence, survey evidence, if properly procured, is well covered by Wigmore's formulation of the rationale that underlies all exceptions to the hearsay rule: "Where circumstances are such that a sincere and accurate statement would naturally be uttered, and no plan of falsification be formed. . . ."<sup>49</sup>

#### SURVEY INTERVIEWEES AS WITNESSES

Some court decisions suggest that a right of the adversary to call survey interviewees as witnesses might help to overcome the hearsay obstacle by fortifying the reliability of the survey evidence.<sup>50</sup> Although in some cases this procedure has proved feasible, and might seem advisable in others, there are strong reasons against making it a general requirement for the acceptance of surveys. Cross-examination of selected survey interviewees is likely to be misleading. It may lend the aura of reliability to an incompetent survey, or it may destroy confidence in a survey which deserves better. Careful studies have shown that, on re-interviewing, one always finds respondents who change their original response.<sup>51</sup> These studies have also shown, however, that such changes, as a rule, do not affect the reliability of the survey. This paradox requires explaining although such explanation might, in fact, encourage the very practice against which it is aimed.

A change of response may be due to any number of causes, *e.g.*, simple response error during either the first or second interview. Even on such factual items as age, a small but definite group of respondents, on re-interviewing, will give a different response. Moreover, changes may occur which reflect opinions and attitudes which, in fact, may have been altered between the two interviews. Finally, the very fact of having been previously interviewed, may precipitate the change. The artificial stimulus of the interview may generate subsequent conversations or inquiries on the part of the respondent which, in turn, will affect the original uncontaminated attitude. A subpoena and what follows may prove even more

<sup>49</sup> 5 Wigmore, Evidence § 1422 (3d ed. 1940).

<sup>50</sup> ". . . nor any of the persons allegedly interviewed were called as witnesses. . . ." General Dry Batteries, Inc. v. Ray-O-Vac Co., 45 Trademark Rep. 588, 594 (1955)—"Any information he [the field representative] could give on the witness stand, would . . . have amounted to hearsay based upon hearsay." Irvin v. State, 66 So. 2d 288, 291 (Fla. 1953). The assurance that the interviewees could be called as witnesses was made a condition of the admission of the survey in *Everitt Hat Company v. Solcum Hat Company*, Milwaukee County, Circuit Court, July 11, 1955 (Wisc.), unreported opinion, quoted in *Barksdale*, supra note 3. "The defendant did not attempt to fortify its survey through any such witnesses." *Oneida v. National Silver Co.*, supra note 46 at 287. Defendant subpoenaed 17 witnesses who had signed the survey. *Quaker Oats Co. v. General Mills, Inc.*, supra note 47 at 431.

<sup>51</sup> Kendall, Conflict and Mood; Factors Affecting Stability of Response (1954).

disturbing; before the cross-examination is finished, so many influences will have become operative since the first interview, that a discovered discrepancy would prove little or nothing. In such a case all the weight would have to be accorded to the first response.

One might argue that there is some value in cross-examining a respondent as to the circumstances of his being questioned and the mode of recording his answers. But these same circumstances can also be elicited from a cross-examination of the interviewers.

The discovery of discrepancies on the individual level need not invalidate the survey results, for while some interviewees may switch from position *a* to *b*, an approximately equal number will switch from position *b* to *a*. The difficulty lies in the fact that if such occasional individual changes are revealed in court through cross-examination, it may tend to impeach the survey because the over-all compensating effect cannot be shown.<sup>52</sup>

There is still another reason against encouraging this practice. Unless the courts protect the survey interviewees from subsequently being called as trial witnesses, bona fide surveys for purposes of settling legal issues will become more and more difficult to conduct. The problem is engendered not so much by the potentially large number of such witnesses, but primarily by a peculiar condition under which surveys operate. Interviewers of private survey organizations are finding open doors and willing respondents because they are scrupulously observing a canon which they share with the Census interviewer. They will not voluntarily identify individual answers by exposing their respondent. This assurance is given either explicitly or implicitly by all reputable survey organizations. Their operations would come to a halt if it were known that an interviewee might have to pay for his cooperation by being called into court and there exposed to the doubtful pleasures of cross-examination. The law specifically prohibits any disclosure of the answers of the respondent in Census surveys in order to insure truthful response and to avoid embarrassment. No other survey response enjoys this protection. A private survey organization cannot assure its interviewees of confidential treatment if the court orders the production of its survey questionnaires.<sup>53</sup> The knowledge of this threat keeps many a survey organization from accepting such legal work.

---

<sup>52</sup> In cases of major importance, the survey organization may attempt to document this by re-interviewing part of the respondents and comparing the paired results.

<sup>53</sup> The issue arose before an FCC examiner concerning the sale of station WGMS. It also arose during the hearings before an examiner for the U.S. Department of Labor concerning a survey of wages paid in the electric lamp industry.

There is a method of avoiding this dilemma.<sup>54</sup> The identification of the survey respondent can be written on a perforated section of the questionnaire and separated after the supervisor has passed it as satisfactory. The separated names will form a record of all respondents, but since none can be identified with any specific response, they can be cross-examined only as to whether they had been interviewed at all. The mutilation of questionnaires, of course, might be misinterpreted, but a proper explanation should eliminate this danger.<sup>55</sup>

For these reasons, then, the cross-examination of the survey respondents should *not* be required as a rule. In the vast majority of cases it will not aid in the evaluation of the survey but might, on the contrary, only confuse the issue. Whether the law, so hesitant to extend the area of privileged communications, will soon respond to these needs is doubtful.

#### THE IMPEACHMENT OF SURVEYS

The discussion thus far has tended to establish that neither the fact that a survey is based on a sample, nor that it relates hearsay evidence should, in itself, bar its admission into evidence. To be sure, there is the danger that the weakness of survey evidence may be hidden under a pretending surface. The following discussion will suggest that under the guidance of expert witnesses any such defects can be satisfactorily exposed and, hence, most danger avoided. The courts generally have refused admittance to surveys in which they have discovered technical flaws, rather than admitting them and permitting their impeachment. Administrative agencies, on the other hand, are more liberal in their practice, partly because they are not bound by the ordinary rules of evidence, and partly as a result of their familiarity with the specific technical problems before them. Moreover, many cases before these agencies are of the sort that can hardly be decided without the assistance of survey evidence.

A court sitting without a jury will seldom hesitate to admit a survey in evidence. The Supreme Court has never either reversed or criticized a trial court for admitting survey evidence in a civil case tried without a

---

<sup>54</sup> It was employed in a survey by the Federal Food and Drug Administration in *United States v. 353 cases \*\*\* Mountain Mineral Valley Water*, Civil No. 565, E.D. Ark., May 1956.

<sup>55</sup> In a survey designed for an FCC hearing, the interviews were concluded with the following statement:

As you may know, reputable survey agencies never make known any individual's opinion without his permission. The sponsor of this survey is seeking an application to operate the service we've been discussing with you. Would you be willing to have your name referred to him in connection with the opinions you have just expressed? This approach is straightforward, but must lead to self-selection of those respondents who are more likely to stand by their outspoken opinions. The result must be a biased group of witnesses of no value to the court.

jury.<sup>56</sup> The court may indeed admit a survey even if it has no confidence in its probative value:

It is doubtful that such an exhibit [a survey of farm machinery outlets in Iowa, conducted by the Statistical Laboratory of Iowa State College] has any relevancy . . . . [B]ut the Court concludes that there is no need to strike the exhibit. It may remain in evidence for what it may be worth. However, it should be stated that the conclusion hereinafter indicated . . . would be the same in absence of this documentary evidence.<sup>57</sup>

When sitting with juries, however, the courts prefer to exclude surveys if, on preliminary examination, they find flaws in them, rather than admit them for whatever they may be worth. Where the line between exclusion and admittance ought to be drawn should depend upon how difficult it is effectively to impeach a bad survey.

### *Improper Universe*

There are three critical points at which a survey operation can fail and provide ground for its impeachment. First, the survey may have been directed at a universe which is irrelevant to the litigated issue. In such cases of obvious error, the court will not need expert advice.

. . . [I]nterviewers stopped [the respondents] in front of one of the appellant's stores in San Francisco and asked them in what manner they spoke of 'Lerner Shops.' Obviously the results of such a survey are of little value in determining what knowledge residents of San Jose had of 'Lerner Shops' . . . .<sup>58</sup>

Or,

. . . [T]he survey, having been limited to retailers, is inadmissible to show that in the market of ultimate consumers the plaintiff's design had acquired a secondary meaning.<sup>59</sup>

### *Inadequate Sample*

It is not always obvious that a survey reflects an improper universe. Sometimes it purports to represent the correct universe but, in fact, does not do so. This is the second point at which a survey may prove inadequate. The universe may be properly selected, but the sample designed to represent it may be faulty. At this point the expert's help, as a rule, will be needed to explain the magnitude of unavoidable flaws or, as the case may be, of any errors in sampling.

The quality of any sampling procedure depends both on its basic design and its execution. The expert will readily discover its deficiencies from the report itself, supplemented by internal documents and such

<sup>56</sup> United States v. United Shoe Machinery Corp., *supra* note 21.

<sup>57</sup> United States v. J. I. Case Co., 101 F. Supp. 856, 868 (D. Minn. 1951).

<sup>58</sup> Lerner Stores v. Lerner, 162 F.2d 160, 162 (9th Cir. 1947).

<sup>59</sup> American Luggage Works v. United States Trunk Co., 158 F. Supp. 50, 52 (D. Mass. 1957).

testimony from the survey staff as may be necessary. Questions directed at discovering the manner in which a respondent was selected for interviewing should provide all the information an expert will need. The survey staff, from the director down to the field interviewers, must be available for cross-examination. While cross-examination of *all* interviewers should be avoided, the court should not refuse to hear as many as are needed to clarify the exact modalities under which the survey was conducted.

To detect deviations from instruction will require a more detailed probing, primarily by questioning the supervisory staff and randomly selected interviewers.<sup>60</sup> Questions should be directed at the institutional safeguards against error (substituting, without permission, respondent *B* for respondent *A*) and against the admittedly rare occurrence of faking parts or the whole of an interview. These safeguards may include proper recruitment, training, and supervision of the field staff, as well as spot controls and double checks of the particular survey sample.

One of the more easily overlooked sampling traps may arise from what is technically called non-response. There are always some individuals in any sample from whom it is impossible to obtain the desired information, either because they could not be located (*e.g.*, were not at home) or because they refused to answer the questions asked of them. An effort to measure the size of broadcast audiences, for example, must go far astray if it bases its findings only upon the people found at home. It will exaggerate the true audience because such people are more likely to listen to broadcasts than those who are away from home and, hence, omitted from the survey. There are several techniques for dealing with this difficulty, all aimed at an estimate of how the non-respondents would have responded had they been reached and interviewed.

As has been pointed out, however, no sample is ever without shortcomings: the exigencies of costs, accidents, and other circumstances may escape control and introduce bias. It is the expert's preeminent task to enlighten the court as to the relevance of such flaws in respect to the measured issues. The point is an important one, for even though a sample have many flaws, it may be judged sufficient for deciding a particular issue. The surveys of the late Dr. Kinsey are illustrative although, of course, they have not come before the courts. The "Human Male" and "Female" were represented only by those ill-assorted men and women who, by one means or another, could be persuaded to be interviewed.

There is a standard method of estimating the survey's true value from

---

<sup>60</sup> It is bad practice to permit the survey organization to present its best interviewer to the court; a survey is only as good as its weakest link.

such an improperly selected sample. First, the degree of under- or over-representation of certain sub-groups is determined, *e.g.*, too many college educated, not enough laborers. One then estimates what the over-all, corrected group average *would* have been had the sub-groups been represented in their true, known proportions.<sup>61</sup> While, of course, such estimates lack the precision of probability samples, they will often be satisfactory if the group measure clearly falls beyond the crucial minimum requirement. But more often, if a sample is improperly designed, the expert will be unable to appraise the size of its bias.

However, even a properly designed and well executed sample may prove wanting, simply because it is too small to provide the desired response, *i.e.*, the sampling error may be too great. This difficulty may derive from the fact that a sample can be sufficiently large to answer some questions, but too small to answer others. Here, again, only the expert can advise the court with precision.<sup>62</sup>

### *Circumstances of the Interview*

The third point at which the validity of a survey must be tested is at its line of questioning and the circumstances under which the interview was conducted. Lawyers know that there is more than one way of posing a question and that the response in each case may be different. Such differences may result from the phrasing of the individual questions, from their sequence, or from the questioning situation. There is a body of experience from which the expert will be able to guide the trier of facts. He will detect bias where the layman sees none, he will know where memory failure will tend to underrate and where vanity may have the opposite effect, and he will know also when, at times, the interviewer's personal opinion affects his respondent's answers.<sup>63</sup>

Yet it is axiomatic, in survey technique, that the danger of question bias increases with the complexity and ambiguity of the questions. If their aim is simple and factual, such as determining the make of the respondent's automobile, neither the form nor the sequence of the questions will make much difference. But in the survey question "As a guess, how

---

<sup>61</sup> An example of this procedure can be found in Zeisel, "Sexual Behavior in the Human Female," 21 U. Chi. L. Rev. 517, 519 (1954).

<sup>62</sup> It might be remembered that the issue of the adequate sample size becomes relevant only with a good sample. As the Literary Digest experience showed, if a sample is wrongly designed its being large makes it only worse, because it makes it less probable that the systematic bias in sampling is cancelled, or at least reduced, by the sampling error working accidentally in the opposite direction.

<sup>63</sup> Hence, the requirement that the interviewer know as little about the purpose of the survey as is compatible with his duties. For a complete examination of all such interviewing problems, compare Hyman, *Interviewing in Social Research* (1954).

much wax would you say there was in glass wax . . . ,” the court rightly found “a built-in bias.”<sup>64</sup>

In another case it was the *sequence* of questions which was found to color the results. At issue was the confusion of “All State” and “All States Life Insurance”:

- (3) What does ‘All State’ mean to you?
- (4) If you wanted All State Insurance where would you go?
- (5) Have you ever heard of All States Life Insurance Company?
- (6) Who would you say owns All States Life Insurance Company?

The court, with justification, criticized the survey for “not fairly presenting the name All State.”<sup>65</sup>

Other aspects of an interview can also become grounds for criticism. Word association tests given to students in a classroom were rejected because their reactions were “bound to differ from that of the buyer in the market place when confronted with the . . . beverage. . . .”<sup>66</sup> As another court remarked, “the issue is not whether the goods would be confused by a casual observer, but [rather] . . . by a prospective purchaser at the time he considered making the purchase. 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.”<sup>67</sup>

A rather subtle source of bias was noted by a court<sup>68</sup> in the selection, as the interviewing area, of “the vicinity of Syracuse, which is not far . . . from the town in which plaintiff’s goods are manufactured,” thus producing an abnormally high degree of confusion with his brand.

Occasionally, the problem at issue will present genuine difficulties to the interviewee. Such an issue arose at a recent hearing before an FCC examiner concerning the sale of radio station WGMS in Washington, D.C. When survey results purporting to show the audience’s preference for “classical” and “semi-classical music” were presented, it was contended that the respondents did not understand the meaning of these terms.<sup>69</sup> The problem of ambiguity might be avoided by posing a *battery* of questions, each of which would cover one facet of the ambiguous concept.<sup>70</sup> Together they would insure that the interviewee has responded to all aspects of the concept.

<sup>64</sup> S. C. Johnson & Son, Inc. v. Gold Seal Co., 230 F.2d 832 (D.C. Cir. 1956), 40 Trademark Rep. 347, 357-58 (1950), aff’d, 90 U.S.P.Q. 373 (1951).

<sup>65</sup> Sears Roebuck & Co. v. All States Life Ins. Co., 246 F.2d 161, 171, 172 (5th Cir. 1957).

<sup>66</sup> Coca-Cola v. Nehi Corp., 27 Del. Ch. 318, 326, 36 A.2d 156 (1944).

<sup>67</sup> Judge Wyzanski in American Luggage Works v. United States Trunk Co., supra note 37 at 53.

<sup>68</sup> Oneida, Ltd. v. National Silver Co., supra note 46 at 286.

<sup>69</sup> 52 Broadcast Telecasting Magazine 68, February 11, 1957.

<sup>70</sup> Compare Lazarsfeld, Comment appended to the Blun and Kalven article, supra note 3.

The hearsay problem also may arise as a result of survey techniques. In *Orbo Theatre Corp. v. Loew's, Inc.*,<sup>71</sup> the trial court refused to admit a survey in which the respondents were asked about their movie going habits *and* those of their families. This added level of hearsay constituted one of the barriers for admittance of the survey.

By way of conclusion, then, we may say that a survey can err in at least three basic ways:

- (1) it may aim at an irrelevant universe;
- (2) although aiming at the right universe, it may not be representative because of faulty sampling, or it may be based on too small a sample and hence render measurements that are not sufficiently precise;
- (3) the mode of questioning, the interviewing situation, or the sequence of questions may tend to reflect inaccurately the characteristics at which the survey aims.

#### EXPERTS AND PROCEDURAL SAFEGUARDS

Courts and juries are presumed capable of judging the value of witness testimony. They are trusted to determine if such testimony is irrelevant or otherwise defective and without weight. Because of a tendency to present survey findings in a way that makes them appear simple and judgeable by the layman, the detection of defects in survey evidence and the evaluation of its weight and significance is not a simple task. Hence, no survey should be presented without accompanying expert testimony. The expert will be needed precisely because of the surface simplicity which almost all surveys, good or bad, tend to display. Thought should also be given to the possibility of appointing impartial survey experts, either by agreement of the litigants or by choice of the court. They would be analogous to impartial medical experts.<sup>72</sup>

Survey evidence has still another peculiarity which should guide its use in court: its production is usually very costly and time-consuming. To confront the adversary with a survey only at the time of trial will almost always constitute an unfair surprise. Surveys, in this respect, ought to be treated much like experiments which, in purpose and mode of analysis, they closely resemble.<sup>73</sup> Two procedural suggestions offered by Professor McCormick in regard to experimental evidence can apply without modification to survey evidence:

---

<sup>71</sup> 156 F. Supp. 770 (D.D.C. 1957).

<sup>72</sup> See Chapter 11, discussing Impartial Medical Experts in Zeisel, Kalven and Buchholz, *Delay in the Court*, (1959); see also Zeisel, "The New York Expert Testimony Project: Some Reflections on Legal Experiments," 8 *Stan. L. Rev.* 730 (1956).

<sup>73</sup> McCormick, *Evidence* § 169 (1954); Zeisel, *Say It With Figures* 132 (4th ed. 1957).

. . . [T]he adversary system . . . must be modified . . . by a rule of court providing first, that no experiment [survey] shall be used in evidence unless reasonable *notice* shall have been given the adversary, with an opportunity to make suggestions as to planning and to be present at the test; and second, empowering the court, in discretion, on application of either party, to appoint an *impartial* person to conduct . . . an experiment [survey].<sup>74</sup>

Where both sides can agree on an impartially conducted survey, this would have the added advantage of dividing the survey costs.<sup>75</sup>

#### TOWARD A SURVEY EXCEPTION

The law with respect to survey evidence is still far from settled doctrine. Thus far, the development has been guided by fears that, since most surveys are hearsay evidence, a bad one might too easily mislead the trier of facts. But our discussion has shown that while the dangers of an uncritically received survey are real enough, they derive not from its hearsay character, but primarily from elements easily opened to expert review. If such expert help is available to the court and the parties to the trial, the dangers arising from the admittance of survey evidence are much smaller than is reflected by the rules which presently govern their admission. These dangers will become negligible if, in the preparation and presentation of survey evidence, the following safeguards are observed:

- (1) All sampling plans, instructions to field workers, questionnaires and other survey instruments ought to be available as evidence of its design.
- (2) The survey staff, from the director down to the ultimate field workers, should be available for questioning as to the survey's manner of execution. The survey interviewees, as a rule, ought not to be required to testify.
- (3) The survey evidence should be presented by an expert witness.
- (4) If a survey is planned during the course of the litigation, the court should explore the possibility of having the survey conducted by stipulation of parties through an agreed-upon or court appointed impartial expert. At that time, such technical requirements as size of sample and other specifications could also be stipulated. If this should not prove feasible, a litigant intending

---

<sup>74</sup> McCormick, Evidence § 169, p. 362 (1954).

<sup>75</sup> . . . [W]here contending parties . . . can agree on the making of a survey (which may or may not mean agreeing to be bound by the *results* of the survey), many if not most of the really controversial problems about survey-making will be eliminated. Waterbury, "Opinion Surveys in Civil Litigation," 44 Trademark Rep. 343, 361 (1954). See also Note, 66 Harv. L. Rev. 498 (1953).

to offer a survey in evidence should be required to notify his opponent early enough to enable him to become an observer in its development. If the survey was completed *prior* to the commencement of the litigation, it should be disclosed to the adversary well in advance of the trial.

If these safeguards are provided, the court should be satisfied that the evidential value of the survey can be appraised objectively; hence, nothing should prevent its admission, provided it is relevant to the litigated issue. If these safeguards are not provided, the court ought to refuse to admit the survey.

While the law might ultimately develop in this direction, its present state gives only small encouragement. A reading of the Model Code of Evidence and the proposed Uniform Rules of Evidence shows quite clearly that the isolated cases in which survey evidence was admitted are far from developing into a settled doctrine. The Uniform Rules do not admit surveys unless they qualify under the state-of-mind exception or, perhaps, as "entries made in the ordinary course of a business," this exception would, at best, apply only to surveys made prior to, and unconnected with, the litigation. However, since surveys provide the best, if not the only, evidence on certain issues, and since expert knowledge in the field has advanced sufficiently to protect the trier of the facts from error, the law may well lower its heavy guard.