

Ethical Aspects of Survey Research

[A report of a Conference held on May 2nd, 1978 at the London School of Hygiene

By STELLA Y. CUNLIFFE and H. GOLDSTEIN

*Committee of Enquiry into
the Engineering Profession*

*Institute of Education
University of London*

This conference, organized jointly by the General Applications and Social Statistics sections of the Royal Statistical Society and the Market Research Society, was a response to an increasing public and professional interest in the personal and social effects of survey work and, in particular, in the issues of confidentiality and privacy.

INFORMED CONSENT

By Roger Jowell (*Social and Community Planning Research*)

The speaker introduced the notion of "informed consent". He outlined its origins in medical experimentation and its eventual codification in the Declaration of Helsinki (World Medical Association, 1964). The idea of informed consent has been translated by social survey workers into a set of criteria to be fulfilled during the conduct of interviews. The difficulty of defining consent, and so recognizing when it has been given, was emphasized. An additional problem facing survey researchers is that there is no pre-existing relationship, such as obtains between doctor and patient, to help overcome communication difficulties or resistance to potentially embarrassing or painful questions. Four criteria are now recognized as necessary for informed consent to be obtained to participation in a survey.

1. *Purpose of survey*

A respondent should understand the general purpose of a survey. This criterion can often be satisfied by a simple explanation at the outset. Sometimes the nature of the survey precludes this. For example, it may be a disadvantage to divulge the true purpose of a survey on racial prejudice. Indeed, various subterfuges may be needed to ascertain the respondents' true attitudes. It has been suggested that respondents be told that the purpose of the survey will be divulged at the end of the interview, but this could well lead to an unacceptably low response rate. A related problem is how a respondent is to be made aware of his position when sensitive questions are asked. For example, information recorded in a survey on unlawful activities of respondents is not privileged in law as is recorded medical information. Technical devices—such as randomized response techniques—have been evolved to preserve the anonymity of this type of survey information, but they are used only rarely and for a limited range of questions.

2. *Sponsor*

The respondent should be told the sponsor, and source of funding, of a survey. This may not cause problems when the sponsor is a reputable academic, or similarly independent, body. Difficulties may arise, however, if the Inland Revenue were to conduct an incomes survey. In market research, the answers to questions in a consumer survey may be biased if the manufacturing sponsor is revealed. A common solution to this problem is to get an independent body to carry out the work, but some might still insist that the original sponsorship be revealed.

3. *Method of selection*

The interviewee should be told how he or she was selected. The most common sampling frame, the electoral register, is well known and easily referred to. It may be argued that use of other lists, such as hospital records or membership lists of societies, breaches confidentiality, and that disclosure of such sampling frames may create resistance on the part of the respondent. This problem is often solved by the list's custodian seeking permission of the selected sample for their inclusion in the survey, but this may well bias the results if many do not reply. Even with this solution there is controversy over whether inclusion in a survey should depend on a positive acceptance or on the absence of a refusal by potential respondents—the latter is the commonest practice.

4. *Access to data*

The respondent should be told how the data will be used and who will have access to them. This criterion is becoming increasingly difficult to implement with the establishment of survey archives. It is closely linked to the first criterion, but has the added difficulty that not all possible uses of data can be seen at the outset. A rigid interpretation of this criterion would require that data were destroyed after the predetermined uses had been exhausted. The existence of survey archives and a flourishing secondary analysis industry indicates that such a rigid interpretation is not widespread and, if it were, it would be very wasteful of valuable data.

It was pointed out that in practice these criteria are only loosely adhered to. Most researchers justify their procedures by the steps they take to preserve the confidentiality of their data and the anonymity of their respondents. Because confidentiality and anonymity seem to be realizable in practice, most effort has gone into achieving them. Often these efforts are considerable but the resultant addition to cost and increased inaccuracy causes practice to vary widely. The conflict between ethical and technical considerations remains.

In conclusion the speaker described an experiment carried out by his own organization in which one group of interviewers in a national survey gave respondents a list of their rights based on the above four criteria with the respondents of the remaining interviewers used as a standard control. The response rates were comparable for the two groups. He felt that this approach might be developed in future codes of practice so that the population became aware of the issues and so became more knowledgeable about our obligations as professional researchers to our subjects.

Discussion

CHRIS WINSTEN (University of Essex) felt that strict codes of conduct were seldom useful in practice, but that the process of discussing them did lead to useful insights and guidelines. He mentioned the American practice which he felt was now generally satisfactory as far as human experimentation was concerned, but pointed out that, for example, the American Psychological Association's code said little about survey research. He was unhappy that the criterion of sponsorship might be weakened by referring in general terms to a grant-giving agency, since such a body would not generally give specific approval to all aspects of a survey. He suggested that other disinterested researchers could be brought into discussions about the propriety of proposed surveys.

The remainder of the discussion showed a very wide range of views. Some felt that surveys should be restricted to gathering a limited range of "factual" information. Others wished for a wide public debate to air political issues about the desirability of particular types of survey. An important issue raised was the availability of information collected from identifiable groups, such as immigrants, which may be used to the detriment of that group. For example, some might refuse to take part in a consumer survey because their responses might be used to exploit them as a member of a particular purchasing group even though it was not used against them as individuals. Research into this topic was thought useful.

PUBLIC USE OF SAMPLED DATA TAPES

By John Boreham (*Central Statistical Office*)

This paper discussed methods of utilizing sample data tapes, especially those produced from the population censuses. One proposal is to release to the public a magnetic tape containing a 1 per cent random sample of the census population, with names and addresses removed. Confidentiality could be protected by safeguards such as "broad banding" of variables like occupation and, possibly, procedures for random "contamination" of the data such as random deletion or addition of variables or cases.

It was emphasized that issues of privacy and confidentiality are fundamentally personal ones but the questions posed about those issues are ultimately political. Proposals for the dissemination of sample tapes (or cards) had been discussed as long ago as 1965. At that time it was stated that the Registrar General had no statutory power to make Census tapes available, but neither was he debarred from doing so. If he did, it was suggested that his actions were unlikely to be legally challenged; this proved to be the case. Since 1965 public attitudes have changed considerably and similar action today by the Registrar General might well be challenged.

In planning the processing of the 1981 Census data, the Office of Population Censuses and Surveys aims to satisfy widespread *ad hoc* requests for tabulations. Mr Boreham felt that it would be more efficient to allow users their own sample tapes than try to meet all requests centrally. An administrative reason is the relative complexity of the organization needed to carry out the work involved. A second reason is that the best method of data analysis for the researcher is to proceed sequentially with decisions on future analyses based on the results of earlier ones. Often this is most satisfactorily done when there is ready access to the raw data.

Lastly, it was pointed out that a magnetic tape containing p variables per observation, with a finite number of categories for each variable, is equivalent to a p -way contingency table. For many sample tapes envisaged the value of p would be small. Problems of confidentiality, etc. involved in the release of such tapes would thus be similar to those of many of the tables currently produced.

Discussion

IVOR CREWE (SSRC Survey Archive, Essex University) agreed that political rather than academic arguments would decide whether sample tapes from the Population Census would be made widely available. The General Household Survey and Family Expenditure Survey tapes were available through the Survey archive; both contained more detailed personal data than the Census. He also referred to U.S. and Canadian experience with the release of 1 and 0.1 per cent sample census tapes for public use. The U.S. Bureau of the Census had simulated user conditions in order to identify particular individuals but, despite valiant efforts, failed to breach confidentiality.

Other discussants emphasized that confidentiality could be ensured by technical means and that the important task facing statisticians inside and outside Government was to inform the public of this and to seek to allay their fears.

DATA PROTECTION AND STATISTICS

By J. Durbin (*London School of Economics and Political Science*)

The speaker thought that the use made of personal information in statistical records was by and large satisfactory as specific steps are generally taken to maintain confidentiality and to guard against the identification of individuals. He believed that the right approach to the regulation of statistical work using personal information was by the negotiation of codes of

practice agreed between the statistical profession and the regulatory authority. The objective in mind in the negotiations should be the achievement of the right balance between the legitimate interests of the statistician, the interests of the individual and the interests of society at large.

Discussion

Mr W. BENJAMIN (Market Research Society) also referred to the balance between public good and private freedom. With a private firm and free competition there is almost a basic assumption of free knowledge, but there can sometimes be a conflict in controlled situations such as arises, for instance, with the Price Commission. Mr E. THOMPSON (Editor of *Social Trends*) again emphasized the major difficulty that the purposes for which data are subsequently used can seldom be anticipated when those data are originally collected. Winding up the discussion, the President of the Royal Statistical Society, Dr H. P. WYNN (Imperial College London), suggested that an independent body should be set up to assess questions of privacy and confidentiality. He also suggested that responsibility for monitoring the proper behaviour by statisticians in these matters lay with the Royal Statistical Society itself.
