Journal of Official Statistics
Vol. 9. No. 2, 1993, pp. 377–381

Discussion
Informed Consent and Notification

Tore Dalenius[1]

1. The Background of the Crisis

In the last two decades, the survey environment has seriously deteriorated. It suffices to point to two features of this development:

1. The growing non-response rate in surveys; in countries where the 1950s and 1960s were known for the public's "survey-mindedness," we now observe what may be described as widespread opposition to surveys.
2. Social researchers are experiencing increasing difficulties in obtaining access to data collected by official statistical agencies. The conference on "Disclosure Limitation Approaches and Data Access" organized by the U.S. National Research Council 1991 is one of many signs of concern.

In my view, the survey statisticians have a moral obligation to society to evaluate the measures introduced to alleviate the serious problem of the public's diminishing confidence in surveys.

The purpose of this note is to direct the attention of government authorities, survey statisticians, and social researchers, among others, to two relevant aspects of today's interviewing practice that should be reconsidered and in my firm opinion changed (or at least modified). These two aspects are discussed in the next two sections. In the third section, I will suggest an approach that would allow social researchers access to the data they need.

2. Informed Consent as a Condition for Data Collection

When considering the pros and cons of informed consent as a condition for collecting data for a survey of individuals, one must not forget that the original domains of application were medical and behavioral research, i.e., domains characterized technically in terms of "stimulus response" or "treatment effect," i.e., applications with some risk of "harm" to the subjects involved. This is clearly not analogous with the situation prevailing in most statistical surveys, whether carried out by an official statistics agency (government surveys for short) or an academic social science organization.

It is my firm conviction that informed consent should not be a condition for data collection in the context of government surveys. What I have just stated must not be interpreted to mean that information about a survey should not be given to the respondents; on the contrary, such information should be provided.

For government surveys, a promise of confidentiality is preferable and sufficient. One reason is that it safeguards fully the

[1] Brown University, Providence, R.I. 02912, U.S.A.

privacy of the subjects; no additional safeguard is needed. And it is far simpler to apply than informed consent. An official statistics agency in a democracy has indeed the power to promise confidentiality and to honor such a promise.

What I have just said is not new. But it is possibly not well understood and not generally appreciated. In support of my opinion, which may appear heretical to some orthodox academicians, especially academicians with no or little experience of data collection (directly from the subjects), I give reference to statements by highly experienced and judicious statisticians. The first reference is to Bryant and Hansen (1975), prepared in the beginning of this era of public concern about invasion of privacy. I remind the readers that Mr. Hansen served for many years as Associate Director at the U.S. Bureau of the Census. The second reference is to Norwood (1991). Dr. Norwood is the former Commissioner of the U.S. Bureau of Labor Statistics. She wrote:

> "We in the statistical system strongly believe that the absolute protection of confidentiality tends to assure the cooperation of respondents in voluntary surveys (and most government surveys are based on voluntary cooperation) and enhances the quality of the responses."

In addition to the principal argument against the necessity of informed consent in government surveys, I present some practical arguments.

1. Informed consent introduces considerable and costly complications of the execution of a survey.
2. Informed consent may prove difficult to use in telephone surveys.
3. Informed consent is not simple to use in the context of surveys for which all or some of the data are collected from administrative records.
4. The provision, typically part of informed consent, that the data collected will be used only for statistical purposes may be interpreted differently by the potential respondents and the survey statistician. The respondents may be likely to interpret this provision to mean that the data will be used only for the subject-matter analysis. The statistician, on the other hand, may include technical uses of the data as parts of the statistical purposes, for example, for editing and estimation.
5. In addition, informed consent may have a negative effect on the response rates in surveys. In recent years, many official statistics agencies have experienced very large increases in the non-response rates in their surveys. It cannot be disregarded that some respondents take advantage of the option not to cooperate (including refraining from completing an interview or diary keeping). In Dalenius (1983), this experience is alluded to by the words "R.S.V.P." in the title.

3. Encouraging Cooperation in Interview Surveys

The non-response rate in interview surveys has long shown a tendency to increase. In many democratic nations – Sweden is one example – the non-response rate in surveys carried out by an official statistical agency is in many important cases well above 15%, to a large extent due to refusals.

Survey statisticians, aware of the potential distortion of the survey results that non-response may introduce, have not remained passive; especially they have devoted great efforts to developing

methods to address the problem. One example is the randomized response technique. Another approach has been an increased use of various kinds of incentives to induce the potential respondents to cooperate. Commonly used incentives are small amounts of money, or some gift. While the use of incentives may be helpful by preventing the non-response rate from becoming devastating, they are, by and large, not sufficient.

In my view, there is need for a radically new measure to counter the potential respondents' inclination not to cooperate. In what follows, I will suggest one measure which may be applied in surveys where the data are collected by interviewers in a face-to-face interview. This measure may be described as an interview which differs in a basic respect from the traditional interview. We will first describe the relevant features of traditional interviewing and thereafter describe in broad outline the new interviewing.

In traditional interviewing, there are two actors:

i. the interviewer; and
ii. the potential respondent; in what follows, this actor will be referred to as a respondent.

The interviewer puts a sequence of questions to the respondent, and records the answers. In a specific case, the respondent may refuse to cooperate, from the very beginning of the interview, or later.

As just described, the interviewing act is one-sided (asymmetrical): data are transferred from the respondent to the interviewer. If we denote the interviewer by "I" and the respondent by "R," the interviewing act may depicted as

$$I \leftarrow R.$$

I will describe the new interviewing along the same lines as those used above.

With regard to the actors, the new situation is similar to the traditional one; i.e., the actors are the interviewer and the respondent.

The key characteristic of the new scenario which differs from the traditional scenario is that the actors exchange data about themselves: the interviewer gives data about him/herself to the respondent, who in turn gives data to the interviewer.

I will elaborate on the data about the interviewer. The data (including name and address) should be recorded on a special card to be handed over to the respondent, who is allowed to keep the card. While informing the respondent about this, the interviewer should mention that if the respondent wants to show the card to a third party (for example, a friend), the respondent should remove the part of the card with the identification particulars beforehand.[2] The interviewer may explain that this is similar to the way the survey questionnaires will be dealt with by the statistical agency. The data should refer to (some of) the variables on the survey questionnaire. Especially, "sensitive" and "personal" data should be recorded on the card. It may prove helpful, if the data are recorded using the same technique as that used on the questionnaire.

On the next page, I present a possible format of a data card, which may be used in surveys which do not ask for responses to highly sensitive questions; hence the card is referred to as a general purpose data card.

[2] If the name and address are recorded on the card, the part with these data may be cut out. Alternatively, the name and address may be written on a label, which is easy to pull off the card.

Identification of the Interviewer:
Name: Tore Dalenius
Address: Providence, RI
Sex:
① male
2 female

Age:
1 less than 30 years
2 30–49 years
③ 50 years or older

Civil Status:
1 not married
② married

Number of Children:
1 no children
② 1–2 children
3 3 or more children

Annual Household Income:
1 less than $25,000
2 $25,000–$39,999
③ $40,000–$74,999
4 $75,000 or more

Housing:
1 lives in an apartment
② lives in a private home

Voting in the Last Election:
① did not vote
2 voted

If you have some additional questions, I will be glad to answer.

The interviewer gives the card to the respondent, explaining the card's content. The respondent is given time to read the card and also ask some questions. Then it is the interviewer's turn to ask the survey questions and record the answers.

The interviewing act may be depicted as follows

$$I \leftrightarrow R$$

where, as before, "I" denotes the interviewer and "R" the respondent.

As to highly sensitive survey questions, the data card should present the answers to one or more of these questions. Especially, this special purpose card may be (part of) the questionnaire used in the survey.

The new interviewing does not preclude the use of the traditional incentives as discussed above.

The new scenario is based on what I consider to be a plausible assumption: that it may serve as a positive incentive. It remains to test this assumption in practice.

Finally, we mention that the new scenario may be adapted to data collection by telephone or by mail. We will not, however, dwell upon this matter.

4. File Splitting as a Basis for Data Sharing

Let F denote a file of records, to which social researchers want to get access. Each record has the format of a data vector with identifying particulars I, classifier data C, and subject matter data $X, Y, \ldots, Z$.

It is not possible to give the researchers access to all data on a record. File splitting may, however, be used to provide classifier data and subject matter data. In what follows, I will give a non-technical description of file splitting; for a technical account, reference is given to Astin and Boruch (1970).

The file is split into three parts:

F_1 comprises a random number R_1, the classifier data C, and the subject matter data $X, Y, \ldots, Z$;

F_2 comprises a second random number R_2 and the identifying particulars I; and

F_3 comprises the two random numbers R_1 and R_2.

The researchers are given access to (a sample or all of) the records accounted for

by F_1. If they experience a problem, they can communicate – by mail or telephone – with the statistical agency, which by means of F_3 would be able to identify the problem record and hence resolve the problem.

5. References

Astin, A.W. and Boruch, R.F. (1970). A "Link" System for Assuring Confidentiality of Research Data in Longitudinal Studies. American Educational Research Journal, 615–624.

Bryant, E.C. and Hansen, M.H. (1975). Invasion of Privacy and Surveys: A Growing Dilemma–Perspectives on Attitude Assessment: Surveys and Their Alternatives. Proceedings of Smithsonian Conference, Santa Fe, New Mexico, April 22–24, 77–86.

Dalenius, T. (1983). Informed Consent or R.S.V.P. In Madow, W.G. and Olkin, I. (Eds). Incomplete Data in Sample Surveys, Vol. 3. San Diego, CA: Academic Press, 85–106.

Norwood, J.L. (1991). Comment. In Duncan, G.T. and Pearson, R.W. Enhancing Access to Microdata While Protecting Confidentiality: Prospects for the Future. Science, Vol. 6, Number 3, 219–239.