Journal of Official Statistics
Vol. 9. No. 2, 1993, pp. 333–334

Discussion
Disclosure Limitation and Data Access

Eleanor Singer[1]

There is an inherent tension between protecting the confidentiality of data and providing access to it.

The Reynolds and Lambert papers address this tension from two distinct perspectives. Reynolds has analyzed some of the generic ethical issues involved, examining privacy both as an absolute right, whose violation constitutes a wrong in itself, and as a condition whose infringement may bring with it varying degrees of injury – for example, public embarrassment, or criminal prosecution. Against these risks he has juxtaposed the benefits of social science research which depends on microdata for analysis. His proposal for reconciling risks and benefits relies largely on procedural safeguards, chief among them a Federal Data Base Review Board which would have the power to authorize research using federal data bases, and to protect such research from compelled disclosure of individual identities.

Lambert, on the other hand, has addressed the technical, rather than procedural, issues involved in protecting the anonymity of individual data against disclosure efforts. And if I understood her paper correctly, the problems of protecting data against disclosure may be far more intractable than we would like to believe.

Both papers, however, make explicit mention of the importance of the public's *perception* of privacy and confidentiality protections, as well as the actual protections themselves. I would like to focus on these perceptions – on the need to look at the issue of privacy and confidentiality from the perspective of members of the public who are potential or actual participants in research. With respect to this group I want to make five points.

First, the public's views may be quite different from those of experts analyzing the actual risks of disclosure, and we need to know what those perceptions are. It seems to me that we also need to reach agreement on what the implications of the statistical analysis are for *actual* risks of disclosure.

Second, what is decisive for participation are the public's perceptions of the risks of disclosure, together with the importance members of the public place on protecting confidentiality. Therefore, along with perceptions of risk, we also need to know, as Lambert suggests, the importance the public places on confidentiality, perhaps in connection with different kinds of data.

Third, it is reasonable to assume that the public's views are not fixed but may be shaped by both long-run influences and short-term events. We know virtually nothing about factors that influence public perceptions of the risks associated with participation in research, and even less about

---

[1] Columbia University, Center for the Social Sciences, School of International Affairs, New York, N.Y. 10027, U.S.A.

what influences the importance attached to confidentiality. It seems plausible that attitudes on these matters are quite changeable, but we do not know that, either.

A little research exists on these questions, but much of it is old, and some has been done in countries other than the U.S. For what it is worth, this research suggests the following: First, people express more concern about confidentiality than they manifest in their behavior, though the reasons for this may be complex. Second, people distinguish between topics, so far as the importance of confidentiality is concerned: they would mind disclosure of some kinds of information much more than they would others, and their preferences in this regard are quite rational. Third, assurances of confidentiality may at times have an effect opposite to what is intended – for certain kinds of content, the more researchers reassure respondents, the more suspicious respondents appear to get.

A fourth point is that policies that fail to address the public's concerns about confidentiality, regardless of how successfully they address the technical and ethical issues involved, will also fail to resolve the basic dilemma. By "basic dilemma" I do not mean the tension between confidentiality and access. I mean, rather, the problem of motivating public participation in research.

What might such policies look like? They might include public information campaigns and other public relations efforts designed to emphasize and illustrate the benefits of research, for example, and designed also to clarify distinctions between administrative and research uses of data. In developing such policies, it may be useful to think of the public as occupying diverse statuses, and therefore as having diverse interests: on the one hand, as respondents with concerns about privacy and confidentiality; and on the other hand as citizens, with a vested interest in accurate and timely data on which to base public planning and policies. It may also be useful to think of concerns about confidentiality as only one of a number of factors potentially affecting cooperation in research.

It ought to be emphasized, once again, that such public relations efforts are not substitutes for ethical and technical safeguards, but are rather intended to supplement them.

Fifth, and finally, I suspect that neither researchers nor administrators ought to be burdened with the entire responsibility for having respondents' interests at heart. As Reynolds recommends, representatives of the public need to be involved in the process of formulating policy and making decisions, and I would argue that this is so even if it imposes additional procedural or substantive demands on the statistical agencies and on researchers.