Sociological Methods & Research
Volume 36 Number 4
May 2008 515-541
© 2008 Sage Publications
10.1177/0049124107313857
http://smr.sagepub.com
hosted at
http://online.sagepub.com

Insiders, Outsiders, and the Editing of Inconsistent Survey Data

Mariano Sana
Louisiana State University
Alexander A. Weinreb
Hebrew University, Jerusalem

Data editing, a crucial task in the data production process, has received little scientific attention. Consequently, there is no consensus among social scientists about how data should be edited or by whom. While some argue that it should be left to data managers and data users, others claim that it is primarily a task for fieldworkers. The authors review these divergent approaches to editing and evaluate the underlying theoretical arguments. Results are reported from a methodological experiment in which different types of actors who are party to the data production and research process were asked to solve artificially generated inconsistencies in real survey data. Results are informative on two counts. First, the least accurate editors were the researchers with no field experience in the survey sites. Second, when provided with only partial information on which to make editing decisions, fieldworkers edited more accurately than both data managers and data users.

Keywords: *data editing; data collection; data inconsistencies; insiders; fieldwork*

Data editing, broadly defined as any action taken to correct errors in the data, is the orphan child of the data production process. In direct contrast to the status of the two other data production tasks, data collection and data processing,[1] it is almost never discussed in the methodological literature. Nor, until very recently, has it received any focused scientific attention (Leahey, Entwisle, and Einaudi 2003, Leahey 2004, Van den Broeck et al. 2005). Consequently, there is no empirical or experimental record from which practitioners can directly infer legitimate data editing practice. Nor is there a general consensus among social scientists about how data should be edited or by whom. Some argue that it should be left to data managers and

data users, though there are considerable differences among the latter about specific steps data editors should take (Leahey et al. 2003). Yet others claim that data editing is a task for the fieldworkers.[2]

In this article, we report the results of a methodological experiment that will push us toward resolving this issue. Specifically, in our experiment different types of actors involved in data collection, data management, and research were asked to solve inconsistencies that we artificially generated in real survey data. The results of the experiment are interesting and informative in two ways. First, they have direct practical applications to the overall process of data production, shining a little more light of "normal science" on the dim terrain of data editing. Second, since our selected actors represent different points along the insider–outsider continuum—a classic continuum in the anthropological and qualitative sociology literature (e.g., Nash 1963; Powdermaker 1966; Agar 1980; Jansen 1980) with some counterparts in the mainstream survey literature (Cicourel 1974; Axinn, Fricke, and Thornton 1991; Weinreb 2006)—they have implications for larger scientific questions about the relative contribution of insiders and outsiders to social science inquiry in general.

Three main sections follow. First, we describe the general scientific context in which data editing occurs and the various approaches to data editing. Afterward, we take a closer look at the question of who should edit inconsistent data, presenting arguments for and against field editing. The third section is empirical: We pose two specific questions and then present the experiment and its results. While we address editing in general, we are most interested in medium- and small-scale survey projects in which the field team is immersed in the research setting and in which data are collected through face-to-face interviews. This is the type of research

Authors' Note: Two awards received by the first author funded the experiment reported in this article: one from the Fund for the Advancement of the Discipline, a small grant program jointly sponsored by the American Sociological Association and the National Science Foundation, and the other from the Faculty Research Grant Program at Louisiana State University. The authors are grateful to the 10 persons who participated in the experiment; to Douglas S. Massey and Jorge Durand for facilitating the experiment's partnership with the Mexican Migration Project; and to Susan Dumais, Jeanne Hurlbert, Erin Leahey, Mark Schafer, Wesley Shrum, Herb Smith, and Susan Watkins for comments on earlier versions of this article. The authors also acknowledge feedback from three anonymous reviewers who contributed to improve the article. Preliminary versions were presented at the 37th World Congress of the International Institute of Sociology, Stockholm, July 5-9, 2005, and the 36th annual meeting of the Southern Demographic Association, Oxford, Mississippi, November, 3-5, 2005. Please address correspondence to Mariano Sana, Louisiana State University, Department of Sociology, 126 Stubbs Hall, Baton Rouge, LA 70803; e-mail: msana@lsu.edu.

project with which, through our own field experience, we have become most familiar.

The Context and Practice of Data Editing

No matter how data are collected—in face-to-face interviews or conversations, telephone interviews, self-administered questionnaires (paper or computer-assisted), and so on—there will be some level of error, including a number of inconsistencies.[3] While some of these will be legitimate, reflecting the fact that people report inconsistent views, others will likely reflect a data collection error. In most settings, for example, a teacher is likely to have more than a primary school education; someone who reports no sexual contact in the last year is not likely to be young, currently married, and coresident with their spouse; a girl of 16 is not likely to be a mother of four. Questionnaires that include these possible contradictions therefore set off alarms. They need to be flagged, and unless our analytic focus is the inconsistency in itself, there need to be procedures in place for dealing with them.

Avoiding the problem altogether by turning inconsistent data into missing values is a poor solution. Generally, the inconsistent data will be partially valid, so a conversion to missing values would waste information. Worse, as inconsistent data are unlikely to happen at random, their elimination would probably generate bias in unknown directions. Automated computer-programmed editing of inconsistencies is also not the answer since, as we argue below, it too can generate bias—albeit in a known direction. In short, data editing of some type is unavoidable. Hence, our interest is who, among all those involved in the production and use of data, should resolve such problems and on what grounds.

Below we review the current state of the art on data editing. We frame our discussion within themes related to chronology, formal survey project norms, researcher incentives, actual field practice, and computer-assisted data collection.

The Chronological Problem

While there is a natural chronological order to data collection, data editing, and data processing, a clear-cut temporal separation between data editing and the other two tasks does not exist. Nor can it exist. In survey projects that involve fieldwork and face-to-face interviews—as noted, the

type of project that we have in mind—data editing usually starts in the field alongside data collection, as questionnaires are reviewed by supervisors or field editors on a daily basis. It then continues into the data processing stage as data entry and consistency checks unveil errors missed (or deliberately ignored) by the field team.

This inability to discretely mark off a period of dedicated data editing that is independent of data collection or data processing lies at the heart of differences about when, where, and how data editing should take place, in particular once the obvious "try harder" step—contacting the respondent again in an attempt to correct the error—has failed (more on that below). Based on our own field experience and our contacts with researchers with and without equivalent background, these differences appear to have given rise to two opposing views.

On one hand are the data collectors. These are researchers with considerable field experience. In terms of their relationship to the respondents and field sites, they can be thought of as "insiders." They tend to believe that the major effort to eliminate inconsistencies should take place as close to the source as possible and soon after the data have been collected. In other words, they think that the major responsibility for solving inconsistencies lies with the field team.

In contrast, there are the data managers and, usually, data users. These tend to have limited field experience and solid formal statistical training. In terms of their relationship to the actual respondents and field sites, they can be classified as "outsiders." They are more likely to believe that researchers and data users should be given the freedom to decide what to do about inconsistent data. In other words, they believe that researchers should receive largely unedited data rather than data whose inconsistencies were solved in the field.

Formal Editing Practices

Given the tension between these insider and outsider views on data editing, it is not surprising to find considerable variation in editing practice from one project to another, especially when we compare U.S. survey projects to those fielded in developing countries.

Outsider editing seems to be commonplace in large U.S. government surveys, with office staff implementing one of two methods. The first is a largely automated system, such as the one used by the Current Population Survey (CPS; U.S. Bureau of the Census 2002). Although CPS protocols allow researchers who would rather use their own imputation methods to

access unedited CPS data, CPS survey staff thoroughly edit inconsistencies along with missing values for more general public use.[4] Alternatively, and representing a more practical solution for somewhat smaller survey endeavors, there is also a manual option. In the National Longitudinal Survey of Youth (NLSY), for example, data editors solve inconsistencies on a case-by-case basis, as long as they feel confident about the edits (Steve McClaskie, NLS User Services, personal communication, December 7, 2006).

In contrast to these outsider methods—one automated and the other manual—a considerable number of research projects conducted by university-based researchers or multilateral agencies, in particular in developing countries, tend to lean toward an insider, field-oriented approach to data editing. The inevitable result is a tension in formal recommendations between the insider and outsider views. In its *Supervisor's and Editor's Manual*, for example, the Demographic and Health Survey (DHS)[5] encourages field teams to edit data errors, especially inconsistencies, while in the field, but only within limits. It is useful to quote at length from these instructions since they signal the difficulty of defining the boundaries of legitimate editing practice.

Where a supervisor or editor identifies major problems, the DHS manual suggests that "it will be necessary to go back to interview the respondent again" (ORC Macro 2002:25). However,

> If a return visit is not possible, try to establish, with the interviewer's assistance, the correct response from other information in the questionnaire. For example, if there is no code circled to indicate a person's sex, you might be able, with the interviewer's help, to determine from the name which sex the person is.
>
> *NOTE: UNDER NO CIRCUMSTANCES SHOULD YOU MAKE UP AN ANSWER*
>
> If it is not possible to return to the household to resolve inconsistencies or missing information, then leave the items as they are. Do not try to fill in every question or to make the questionnaire consistent. (ORC Macro 2002:25)

This is the DHS's formal statement of practice. The careful wording cannot disguise the tension—or in stronger terms, the hint of contradiction—in the instructions. Fieldworkers are told to edit, but not by making up an answer.

Three pages later there is an even sharper example of this tension, this time in relation to the editing of the respondent's age, a crucial variable in the DHS data as it determines the very eligibility of a respondent for the survey as well as the validity of estimated age-specific rates that are one

of the DHS's key functions. Here, DHS field editors are given direct instructions to infer a given respondent's age from specific variables:

> If at all possible, an effort should be made to revisit the respondent to resolve inconsistencies with age since it is one of the most important pieces of information collected in the questionnaire. If a revisit cannot be scheduled, it may be necessary to look at other information in the Household and Women's Questionnaires in an effort to resolve the inconsistency. Items that should be considered include the following:
>
> - Age recorded for respondent in Household Questionnaire
> - Number of live births
> - Date of birth of respondent's first child
> - Date or age at first marriage. (ORC Macro 2002:28)

Researcher Incentives

To the extent that major research projects such as the DHS formally embrace some level of field editing, they provide procedural legitimacy (M. Suchman 1995:579-80) to something that actually happens in the field. That is, field teams do edit data to some extent, either openly and following protocol (as in the DHS) or else covertly. To understand why, one needs to consider that all parties involved in the production of data, from the principal investigators to the fieldworkers, are under institutional pressure to produce high-quality data and to do so on schedule and within budget (Prewitt 1983). Ideally, high-quality data are those with high levels of validity. However, given the difficulty of validating data, all those involved in data production tend to focus on more visible markers of data quality. Foremost among these are whether data are complete and consistent. Given a target sample size, completion and consistency are the markers by which data producers are held accountable.

More specific to the data-producing field staff, it is useful to distinguish between different sets of incentives. First, both interviewers and their supervisors are institutionally propelled toward editing—even when they may not have sufficient input to do it confidently—since resolving errors exonerates them from the failure to collect correct data in the first place. It also allows them to maintain a certain level of productivity—that is, to produce a certain number of complete, consistent questionnaires in a given time frame—which is at least informally used as a measure of their respective abilities. Second, supervisors have an additional incentive to correct problematic questionnaires. To carry out their assignment, they

need to maintain a comfortable working relationship with their subordinates. Refusing to fix interviewers' errors does not service this goal since it implies a degree of callousness toward the interviewers or lack of concern with data quality, respectively lowering the team's morale and productivity.

Field Editing in Practice

It is useful to describe what field editing looks like. We draw on our experience as former field directors of the Latin American Migration Project and the Malawi Diffusion and Ideational Change Project[6] and on informal conversations with researchers in other developing countries.

Once an interviewer completes an interview and turns in the questionnaire to her or his supervisor, a review process begins. Either specialized field editors or field supervisors are in charge of this task. When they detect errors, the nearly universal recommendation is to use "callbacks."[7] That is, the supervisor sends the interviewer back to the respondent to fill in the missing values or resolve the inconsistencies with real rather than imputed data.

Callbacks, however, are not problem-free: Budgetary constraints and deadlines may preclude additional visits or limit their number; when the interviewers do go back, the respondents may be out, away, busy, unwilling to open the door again for the interviewer, or either unwilling or unable to report a more complete or consistent set of responses. If any of these occur, there are two courses of action open to fieldworkers. They can leave the resolution of all remaining data problems to the outsiders—the data managers and data users—or they can resolve the problems themselves. The first of these is the customary practice in most surveys in the United States;[8] the second one, however, is the most consistent with fieldworkers' own incentives.

What does this field resolution of an inconsistency look like in practice? Typical accounts begin with a field supervisor checking a given questionnaire and noticing an error. The supervisor then meets with the corresponding interviewer to discuss the problem. Sometimes the interviewer reports to have asked the question, to remember the answer, but to have forgotten to fill in the corresponding form or to have miscoded the response. In other cases the interviewer describes how, when the question was asked, the respondent's attention shifted to something else (e.g., the baby started to cry and the respondent got distracted), which led to a strange response. Often, based on his or her recollection of the respondent,

the interview, and the context in which the interview took place, the interviewer can gather enough confidence to produce a reasonable suggestion on how to correct the data error. The supervisor can also contribute to this process, either by positing helpful questions to the interviewer or by considering the specific problem in light of general patterns that he or she has observed in the survey setting. Either way, the exchange can result in the resolution of the problem.

Computer-Assisted Data Collection

The discussion thus far has largely focused on traditional, paper-based questionnaires. The debates between insider and outsider approaches to editing are equally appropriate to computer-assisted interviewing, which, notwithstanding a number of practical limitations, is likely to become the dominant practice in the near- to midterm in both developed and developing countries (for a discussion of types, benefits, and drawbacks, see Greene 2001; De Leeuw, Hox, and Kef 2003; Fletcher et al. 2003; Mensch, Hewett, and Erulkar 2003; Ice 2004). This transition from paper to screen is likely to reduce the number of inconsistencies that are generated in any given interview since built-in checks can prompt the interviewer to probe the respondent when an answer to a question is inconsistent with the answer to a prior question or is beyond a minimum or maximum threshold value. Yet some inconsistencies will remain. The main reason (for some others, see U.S. Bureau of the Census 2002:7-8) is that such checks must be flexible enough to allow for respondents who are unwilling (or unable) to respond consistently. In addition, they must not be too frequent or time-consuming so that they unnecessarily irritate respondents. In either case, too many probes or too inflexible an automated editing program may increase the risk of premature termination of the interview. Overall, data collected via computer-assisted interviewing likely attenuates the prevalence of inconsistencies, but it does not remove them entirely.[9]

Two Opposing Views of Field Editing

When we first described, at a professional conference, how field editing works, a senior methodologist stated that he was "horrified" by the idea of fieldworkers editing the data in any way. His aversion was consistent with findings by Leahey et al. (2003), who requested advice from a sample of social scientists on what to do about specific inconsistencies—presented in

an experimental vignette—in survey interviews carried out by a hypothetical researcher. Three quarters of their subjects objected to the edit proposed in the vignette, as did as many as 90% of a different sample of scholars in a subsequent experiment reported by Leahey (2004).

Discomfort with field editing can be articulated as three related arguments. They address the problem of legitimacy, the presumed advantage of training or educational credentials, and the risks that derive from fieldworkers' preconceptions. Not surprisingly, there are also counterarguments. In this section we briefly review each of these.

Contra Field Editing

The first argument against field editing questions its scientific legitimacy. There is no discussion of field editing, let alone approval, in the methodological literature. It is not generally recorded—marks may be made on the actual questionnaire, but they do not make it into the data files. Nor are there standardized criteria for the task. Rather, edits are customized to each individual case, and different editors might solve the same inconsistencies differently. As a consequence, field editing is not easily amenable to procedural replication, a key feature of scientific research. By contrast, automated edits embedded in computer code seem to be a better choice.

The credentialist argument builds on the assumption that educational attainment, in particular higher education, is an indicator of the type of intelligence needed to resolve inconsistencies or that there are specific types of statistical issues that data editors need to be aware of to resolve inconsistencies. Since interviewers, particularly in developing countries, can rarely claim a college degree and since they have little to no statistical training or research experience outside of data collection, they should not be allowed to edit the data. Rather, such decisions should be left to more educated outsiders, in particular those with higher education and advanced statistical training.

The preconceptions argument is concerned with bias that field editors can introduce into the data. Specifically, fieldworkers may have some prior idea about relationships that are at the core of the investigative enterprise—or that principal investigators, their employers, would like to find—and they will rely on these ideas to resolve data problems (as coders of medical records in a psychiatric clinic did in Garfinkel's [1967] early ethnomethodological experiments). As a result, reported behaviors or attitudes of core interest to the lead investigators may be artificially exaggerated or suppressed, distorting analysis in predictable ways (Blalock 1961; Zeller and Carmines 1980).

Pro Field Editing

Arguments in favor of field editing address the same issues, but from the perspective of data collection insiders. First, although a priori the scientific legitimacy of field editing appears to be low, the question remains whether data editing at other stages of the research process is any better. It is certainly not proven, for example, that manual editing by outsiders (e.g., the NLSY model) generates more valid or reliable data than field editing. Somewhat surprisingly, one can be equally skeptical of automated editing processes (e.g., the CPS model). Although it is true that they are inherently more standardized and therefore more reliable than field editing, this does not address whether reliability, rather than validity, should be the primary goal of data editing. For a simple example, assume an editing rule that says "if woman = pregnant and virgin = yes, then virgin = no." This edit will be incorrect every time that a virgin was mistakenly recorded as pregnant, but it will never be wrong the other way around. In other words, standardization prioritizes reliability but in so doing risks bias, thereby threatening validity.

Of course, automated edits need not be so simplistic. But the more complex they become—the possibilities for adding conditional statements to the code are endless—the more intensive the programming task is, the more prone to programmer error, and the more it resembles manual editing. Moreover, not all projects can afford this complexity. Instead, most use computer code to *detect*, rather than to *fix*, inconsistencies.

In short, the outsider approach certainly looks more scientific than its insider alternative. It is conducted by scientists sitting in front of computer screens rather than fieldworkers hunched over clipboards. Since it is based on standardized algorithms embedded in computer code, it can be replicated. However, since the relative performances of the different methods appear never to have been empirically tested or verified, it is not clear that outsiders' preference for office-based editing—whether automated or manual—is anything other than a disciplinary norm or an artifact of our tendency to fetishize science and technology. Indeed, this comes across quite strongly in Leahey (forthcoming). She interviewed institutional review board chairpersons, National Institutes of Health and National Science Foundation program officers, and sociology journal editors on the problem of data editing. Most implied that they trust researchers to do their job well and that scientists gain their trust by following the technical norms of the scientific inquiry. In other words, "Fulfillment of the technical norms is

taken as satisfactory technical competence and establishes trustworthiness in that respect" (Barber 1990:138). This is little different from saying that data editing by outsider researchers has scientific legitimacy because outsider researchers are scientists and what scientists do is scientific. In more classical sociological terms, data editing by scientists is legitimized on rationality grounds supported by tradition (Weber [1922] 1968).

The credentialist argument rests on equally questionable assumptions. While higher educational attainment signals presumed ability to solve complex problems, it is not clear that this is the skill required to resolve an inconsistency. An intersection of three literatures from across the social sciences suggests that it is not. The first is largely anthropological. It focuses on the value of "situated knowledge" (Haraway 1991) and researchers' insiderness (e.g., Agar 1980; Stocking 1983) in facilitating an instinct for sniffing out inaccuracies in the field and detecting what inconsistent piece of information is most likely to be wrong in a particular setting. The second is microsociological. Since Garfinkel's (1967) ethnomethodological experiments, it has been clear that researchers at all levels—like social interactants in general—draw on "contextual knowledge" to code or assign meaning to given phenomena (e.g., Heritage 1984; see also microsociological extensions in Collins 2004). This is equivalent to the situated or insider knowledge in the anthropological literature since it refers to how our awareness of context-specific sounds and signs, verbal and otherwise, allow us to accurately read interactants' intents. The third relevant literature comes from psychology and is concerned with the nature of decision making under pressure or with limited information (for an engaging compilation, see Gladwell 2005). The gist of this literature is that theoretical knowledge and even large quantities of presumably relevant data offer little guarantee of achieving success in decision making (Schwartz 2004). On the contrary, the key to excellent decision making is more a function of one's ability to pinpoint the few pieces of truly useful input among large quantities of presumably relevant but distracting information. Rather than being an automatic byproduct of higher education, skilled decision making depends most on task- and context-specific instincts, which are in turn enhanced by task- and context-specific experience.

Applied to the question at hand, these three literatures suggest that higher education credentials and specialist statistical skills may make no particular contribution to the resolution of inconsistencies in survey data. More important is the extent to which the fixer of inconsistencies is an insider, familiar with the cultural setting in which the data were collected.

Fieldworkers clearly are. If they are asked to fix data from an interview that they themselves conducted, they can draw on additional information that is rarely made part of an official data record. Examples include the general feel of a respondent's house, the relationship among family members and neighbors, side comments about a range of topics outside the interview's main focus, comfort level with the respondent, minor events during the interview that were distracting to the respondent and/or the interviewer, and so on. Field supervisors, in turn, can draw on a similar set of markers, albeit more general ones, since they are based not on the specific interaction but on having been in the area, talked informally to local residents, and consequently developed ideas and instincts about what things look in and out of place in that setting. Additionally, supervisors may have other types of useful information that would affect the resolution of inconsistencies, such as impressions about the reliability of given interviewers, knowledge of clashes between team members, and so on. In short, it seems reasonable to assume that fieldworkers can draw on a range of relevant task- and context-specific information, provided by their position as insiders, when resolving inconsistencies, while at the same time efficiently discarding seemingly related pieces of information that are not needed for accurate editing. Having spent little or no time in the field, this skill will be foreign to outsiders, irrespective of their academic credentials or specialist statistical training.

Finally, the preconceptions argument against editing by fieldworkers is also weaker than it initially appears since it ignores the powerful incentives that affect the researcher as a special type of organizational actor. A less disingenuous portrayal recognizes that all parties involved in research, starting with the researchers, have preconceptions and are subject to incentives. The question is how much these affect different actors' readiness to fix data problems in ways that will serve their own analytic ends. A recent flurry of high-profile cases shows that data fixing is a temptation to which some researchers succumb.[10] Still more worrying is the fact that unethical data manipulation seems to be emerging—or perhaps is merely increasingly acknowledged—as a generalized institutional problem within science. Of a sample of 3,247 U.S.-based scientists, for example, Martinson, Anderson, and de Vries (2005) report that 15.3 percent admitted "dropping observations or data points from analyses based on a gut feeling that they were inaccurate," 6 percent admitted "failing to present data that contradicted one's own previous research," and a variably significant proportion of scholars admitted to skirting other types of

research ethics. Rubenstein (2006:3) describes one of the most common temptations:

> In economics, there is no tradition of checking data and repeating experiments. In the few cases in which I conducted experimental research, I myself felt the pressure not to search further at a stage in which the experimental results went in my favor and to check findings seven times when they appeared not to support the assumptions I was sure were correct.

The manipulation of data (or of models) is a predictable result of pressures to produce publishable work. We see no reason why this temptation should be any different for the outsider social scientists charged with data editing. By contrast, being less invested in the analytic results of the research inquiry—their job ends with the collection of data—fieldworkers have no incentive to push the data in any particular direction.

Who is Best Suited to Solve Inconsistencies?

Given that inconsistencies are ubiquitous in survey data, that some will remain even after repeated attempts to obtain correct information, and that solving them inevitably involves some level of inference from extant information, a sensible question arises: Who is best suited to solve inconsistencies? Without any empirical backing, the opposite actors in our drama have already answered this question for themselves. Fieldworkers are best suited, say the insiders; not so, say outsiders—data users or office-based personnel must do it.

Our first question pits these two groups against each other. The case for field editing notably relies on situated knowledge and field experience, which we have subsumed under the term insiderness, as the key attribute of good editors. The question we pose therefore is "Are insiders better editors than outsiders?"

The second question is "Does editing ability improve with additional input?" This question is important because, by virtue of their insiderness, fieldworkers have more information, and presumably a better sense of which pieces of information are really relevant, than the outsider office workers. It also has an important resource-efficiency element if, for example, fieldworkers need less time than outsiders to gather and cognitively sort the necessary input to resolve the inconsistency—which is likely to be the case as field editing takes place shortly after the interview.

To address these two questions, we set up an experiment.

Experimental Design

Ideally, researchers would evaluate field editing as it occurs. A researcher would take note of the data problems found by the supervisor, ask the field team to solve them, and either go personally or send a research assistant to the respondents' households to find out the real solutions to the problems. Over time, the researcher would amass a large number of these episodes to produce a rigorous analysis of field editing techniques.

Such an approach would be quite difficult to implement in practice.[11] Consequently, we evaluated edits made by different types of survey and research actors in an experimental setting in which fieldworkers' solutions could be considered as an approximation of their editing ability in the field. We conducted the experiment with the cooperation of the personnel of the Mexican Migration Project (MMP). The MMP is a binational survey research project based at Princeton University and the University of Guadalajara. Every year, it conducts a set of migration surveys in selected Mexican communities. A team of four or five interviewers and a supervisor carries out each MMP survey, typically 200 households, using an "ethnosurvey" instrument (Massey 1987). The interview is focused on migration to the United States and collects a wide range of individual- and household-level information as well as complete life histories of the household head and the household head's spouse.

We measured the relative editing ability of various actors located in different positions of the data production and research process. To evaluate how good or bad edits were, we needed to know the reported values of the edited data. Consequently, we designed the experiment using data already collected. After selecting a sample of questionnaires, we altered selected pieces of information to force inconsistencies into the data and asked the participants to infer which piece of information was wrong. In each case, the participants browsed through a real MMP questionnaire on a computer screen, with true data except for the alteration in question. All questionnaires used in the experiment were randomly selected from four communities surveyed between 2000 and 2002, subject to the condition that the respondent (the household head) had migrant experience in the United States.

By altering selected pieces of information[12] we created three types of inconsistencies:

1. Inconsistency between migration history and labor history—for example, the migration history indicates that the respondent was in California in a particular year, but the labor history lists the respondent's job, for that entire year, in Mexico.
2. Inconsistency between migration history and marital status—for example, the respondent appears as married at the time of migrating, but his marital history shows no marriage before migration.
3. Inconsistency concerning documentation status in the United States—that is, two different parts of the questionnaire disagree on the respondent's documentation status during a migration trip.

In each case, only one component of the pair was altered, leaving the other intact. We asked the participants to identify which component from each pair was incorrect. For the measurement of editing ability of different participants or groups of participants, we simply computed percentages of correct solutions and their respective confidence intervals.

To address the hypothesis that proximity to the field improves editing ability, we recruited 10 participants representing four different positions in the data production and research process: 4 interviewers, 2 field supervisors, 2 data managers, and 2 data users. Each of the four positions indexed a different level of insiderness. The interviewers, all but one of whom were undergraduate students at the University of Guadalajara, had a number of years of experience doing fieldwork with the MMP. They therefore represented the maximum level of insiderness. The field supervisors, both with master's degrees, had been in charge of fieldwork in multiple MMP survey settings. With less direct contact with respondents, they represent a slightly lower degree of insiderness. All interviewers and field supervisors were based in Mexico. The rest of the participants were based in the United States. The 2 data managers included a senior project assistant and the MMP project manager at the time of the experiment, both holders of master's degrees. The 2 data users, both with PhDs, had vast experience using the MMP data for research. Neither the two data managers nor the two data users had any field experience. However, as the data managers were more involved in the nuts and bolts of the data production process, we rank them higher than data users on the insider–outsider continuum.[13]

Our second empirical question relates to the effect of additional information on editing accuracy. To answer this question, participants were asked to edit each inconsistency twice. As input for their initial edits, which we refer to as "restricted," they were given only a fraction of the information in the questionnaire. After making this first edit, we allowed them to review all

existing information: the full record for the migrant in question, the data corresponding to other members of the household, and even the questionnaires of any other households from the same community. They then produced a second edit, which we call "unrestricted."[14]

We selected 20 questionnaires from each of the four communities and created one inconsistency in each questionnaire. The 20 inconsistencies broke down into 7 concerning migration and labor history, 7 involving migration and marital history, and 6 related to documentation status. With four communities and 20 edits per community, each participant had to edit 80 items. Each edit was made twice, with restricted and unrestricted information, for a total of 160 edits per participant and a grand total of 1,600 edits.

Edits were produced in the same order by all participants, completing all edits for one community before moving on to the next. Before beginning each community, the participants were given a description of the community in question. Not all Mexico-based participants had worked in all the selected communities, but those who did not had fieldwork experience in similar settings. For the U.S.-based participants, this information gave them their first and only introduction to each community, resembling the typical basic information that is usually available to any user of a survey data set. To avoid contamination effects, each participant worked on a separate computer, and participants were not allowed to discuss their edits.

An obvious concern was whether the pieces of information selected for the experiment might have been edited themselves. We could not know this with certainty, but because the parts of the questionnaire that we selected for the experiment are only a very small fraction of the information collected, we deem this likelihood low. A related concern was the possibility that a participant had to edit part of an interview that he or she had personally conducted. Again, we did not consider this a problem since the probability of matching an interview with the fieldworker who actually conducted it was less than 10 percent, and in any case, the interviews and the experiment were at least 2 full years apart, and the fieldworkers in the experiment had interviewed quite a large number of households in those 2 years.

We do acknowledge, however, other limitations in the study design. First, the experiment denies fieldworkers the possibility to use memories of specific interviews to address problems with those interviews. We think that this is to the detriment of interviewers' editing ability since the recency of the interview, and memories from it, constitute an important advantage that they have over participants in all other positions. As a result, the experiment places the interviewers in largely the same position as the supervisors, asking both to resolve problems with *other* people's interviews using three

sources of information: restricted/unrestricted material from that particular interview, memories of past interviews in similar settings, and general knowledge about the research area. Consequently, we think that the experiment underestimates the advantages of field editing for interviewers and, to a lesser extent, for supervisors.

Second, the number of participants recruited for the experiment is relatively small—a consequence of this being a pilot study and of the objective difficulty in recruiting a larger number of actors involved in the same data production process in addition to users of the data in question. As we show, however, the limited number of participants did not prevent us from reaching some suggestive conclusions.

Finally, the experiment was intended to measure editing ability of individuals in different positions of the data production process, but in real editing practice collaboration between these individuals can, and often does, take place. An interviewer with supervisor input might solve an inconsistency more accurately than the interviewer alone, or office workers might ask the field team for their input when solving inconsistencies found in the office. Our experiment does not address collaboration across positions. Instead, it focuses on the ultimate question of who among these actors is best suited to do the job.

Results

Our reporting of results stresses comparisons across positions. However, considering that with the exception of the four interviewers, we have only two participants in each position, it is fair to ask whether any apparent difference in performance between two positions is really the result of just one participant being a gifted, or a particularly weak, editor. To address this and other questions related to individual performance, Table 1 presents all participant-specific scores.

No single interviewer appears to have consistently over- or underperformed all of his or her peers, although Interviewers 1 and 2 outperformed Interviewers 3 and 4 on all inconsistencies combined, both restricted and unrestricted. In addition, Supervisor 1 tended to score higher than Supervisor 2 by several percentage points, and there was an even larger gap between the two data managers. Likewise, one of the data users produced more accurate edits than the other data user. Of special importance for the analysis that follows is the fact that even though Data User 1 tended to be a better editor than Data User 2, both of them were generally worse editors than everyone else. In other words, the differences that we report below between data users

Table 1
Percentage of Correct Guesses by Participant and Type of Inconsistency

	Participant									
	Fieldworkers						Office Workers			
	Interviewers				Supervisors		Data Managers		Data Users	
Type of Guess	Int 1	Int 2	Int 3	Int 4	Sup 1	Sup 2	DM 1	DM 2	DU 1	DU 2
Restricted										
Migration history/labor history	67.9	67.9	53.6	60.7	64.3	46.4	64.3	39.3	28.6	32.1
Migration history/marital history	57.1	82.1	67.9	78.6	85.7	67.9	89.3	50.0	60.7	35.7
Documentation	95.8	79.2	62.5	66.7	70.8	79.2	58.3	58.3	41.7	29.2
All three combined	72.5	76.3	61.3	68.8	73.8	63.8	71.3	48.8	43.8	32.5
Unrestricted										
Migration history/labor history	78.6	89.3	75.0	75.0	78.6	78.6	82.1	64.3	35.7	39.3
Migration history/marital history	82.1	89.3	85.7	67.9	89.3	75.0	96.4	71.4	71.4	71.4
Documentation	95.8	70.8	75.0	62.5	75.0	70.8	70.8	75.0	62.5	58.3
All three combined	85.0	83.8	78.8	68.8	81.3	75.0	83.8	70.0	56.3	56.3

Table 2
Individual Guessing Matches Between Fieldworkers and Office Workers, on All Inconsistencies Combined

Type of Guess	Office Team				
Field Team	Data Manager 1	Data Manager 2	Data User 1	Data User 2	Total
Restricted					
Interviewer 1	1	1	1	1	4
Interviewer 2	1	1	1	1	4
Interviewer 3	0	1	1	1	3
Interviewer 4	0	1	1	1	3
Supervisor 1	1	1	1	1	4
Supervisor 2	0	1	1	1	3
Unrestricted					
Interviewer 1	1	1	1	1	4
Interviewer 2	=	1	1	1	3
Interviewer 3	0	1	1	1	3
Interviewer 4	0	0	1	1	2
Supervisor 1	0	1	1	1	3
Supervisor 2	0	1	1	1	3
Total	4	11	12	12	39

Note: A number 1 indicates that the fieldworker (row header) edited better, overall, than the office worker (column header) on the type of edit in question (restricted or unrestricted). A number 0 indicates the opposite. The equal sign shows a tie between Interviewer 2 and Data Manager 1 on unrestricted guesses.

and other participants are not the result of one data user being a particularly substandard editor.

Additional reassurance for our focus on participants' positions is provided by kappa statistics that measure the chance-adjusted level of agreement between editors. Results (available from the authors) show that the degree of within-group agreement (among all four interviewers and between the two supervisors, the two data managers, and the two data users) tended to be higher than the degree of across-group agreement. Specifically, across four positions, three types of inconsistencies, and two types of edits (restricted and unrestricted), within-group kappa statistics were higher than the average kappa coefficient for all across-position pairs of participants in 18 of the 24 cases. The implication is consistent with our main premise that position in the data collection and production process matters: Editors in the same position are more likely to think alike when it comes to solving inconsistencies than editors in different positions.

Figure 1
All Inconsistencies Combined: Percentage of Correct Guesses and 95 Percent Confidence Interval, by Position, Restricted and Unrestricted

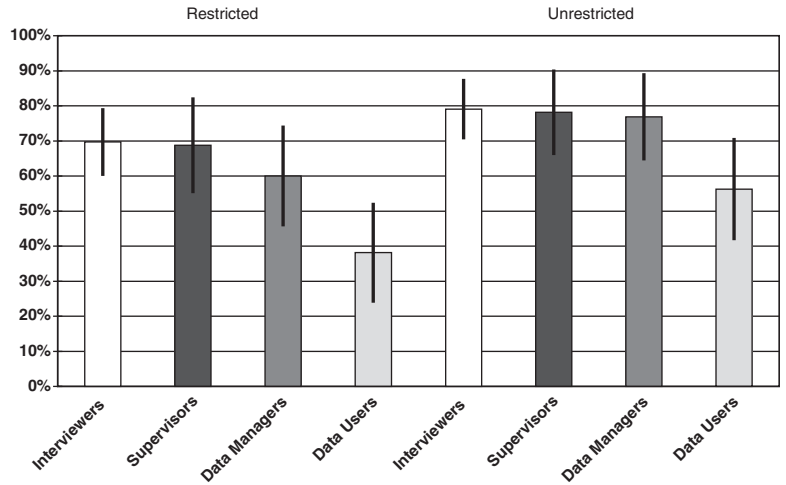

Table 2, constructed with the results presented in Table 1 for edits on all three types of inconsistencies combined, shows the results of the 48 individual editing matches between members of the field team and the office participants, who lacked field experience. A number 1 indicates that the fieldworker edited more accurately than the office worker; a 0 signals the opposite. For example, the number 1 on the top-left corner shows that Interviewer 1 produced overall more accurate restricted edits than Data Manager 1. These scores show that fieldworkers prevailed 39 times, with 1 tie and 8 defeats. Both data users were soundly outperformed by the fieldworkers' better editing ability in all matches. Data Manager 2 was outperformed in 11 out of 12 matches. Only Data Manager 1 scored a positive count against fieldworkers but was still outperformed in 4 matches and was tied once.

The percentages of correct edits and their corresponding 95 percent confidence intervals, by position and for all three types of inconsistencies combined, are presented in Figure 1. Restricted edits are shown on the left and unrestricted edits on the right.

With restricted information, interviewers and supervisors (editing accuracy of about 70 percent) edited significantly better than data users, who

in fact edited inaccurately most of the time (accuracy below 40 percent). Fieldworkers also produced better edits than data managers (the latter's accuracy was about 60 percent), although the difference was not statistically significant. Each position improved their edits when complete information was available. With unrestricted information, data managers positioned themselves at the same level as interviewers and field supervisors, with all of them reaching a 75-80 percent accuracy rate. Data users still lagged behind after complete information was made available. With an editing accuracy rate slightly above 55 percent, they were clearly the worst editors.

Discussion and Implications

Inconsistencies are inevitable in survey data, and in the absence of formal methods, there is no other resource than human judgment to solve them. In our experiment we evaluated, for the first time, the relative abilities of those who produce and use survey data to edit inconsistencies. The good news is that given complete information, those most involved in the data production process fixed this kind of problem accurately more than 75% of the time. The bad news, at least for some, and notwithstanding our small sample size, is that researchers not involved in the data collection process or in other aspects of data processing appeared to be much less skilled at solving inconsistencies. Their editing accuracy did not reach 60%, even with complete information.

In answer to the first analytic question that we posed earlier, we have therefore provided some evidence that in the world of data editing, the editor's position on the insider–outsider continuum matters. Data users, furthest removed from the dirt of data collection and from the field sites more generally, were the least accurate editors. They were both outperformed by all six fieldworkers despite their experience and familiarity with the data set. In addition, with restricted information, fieldworkers edited more accurately than the less involved data managers. These two results are informative, and they advance our understanding of the scientific problem of data editing, particularly given the fact that the experimental design, as discussed above, diluted the editing advantage of fieldworkers.

Concerning our second question, all groups improved the accuracy of their edits when provided with additional information. Data managers, in particular, were able to close the gap on fieldworkers when they had full information. We think that this result lends yet more support to field editing, which in actual practice benefits from the recency of the interviewing

experience and all the related information that would not be available to data users later or that would be costly for them to gather.

These results may seem modest and commonsensical to those who engage in fieldwork. Yet as we argued above—and as shown by prior research—most researchers feel uncomfortable with the idea of unvalidated and unreplicable data editing, and many firmly believe that fieldworkers should not edit data. Our results question the validity of these feelings and beliefs. Admittedly, we see some risk in generalizing too readily from these results. In addition to sample size limitations, the diversity across research teams and the conditions in which they labor make it difficult to define a population of research teams onto whom we could legitimately generalize. Nevertheless, we believe that our results have at least two practical implications.

First, the resolution of inconsistencies should not be left to the researchers. In other words, notwithstanding their professional status and credentials, data users who do not go to the field appear to be the weakest link in the chain of potential editors. Second, fieldworkers appear to be the most efficient choice for accurately resolving inconsistencies since, by virtue of situated knowledge, experience, and even their mere presence in the survey site, they naturally have a larger stock of relevant information than office workers. In addition, the cost of fieldworkers' time is usually far lower than the cost of office-based data editors.

We do not, of course, expect all researchers to adopt these practices overnight. Based on a long series of informal conversations with colleagues, however, we believe that many in the social science community who spend time in the field have already made moves in this direction, albeit informally and under the radar or else, primed by prior experience, are ready to do so in the future. This suggests that field editing will become increasingly acknowledged and legitimate. It is instructive to note that this will not be the first time that a method of data production initially dismissed as unstandardized and unscientific will have become accepted in the methodological mainstream. One recent precedent, for example, is the move toward more conversational styles of interviewing (Dijkstra 1987; Suchman and Jordan 1990; Beatty 1995; Schober and Conrad 1997).

Either way, we think that the uniformity of the differences between field staff and data users in our experiment speaks for itself. Of all the players in this particular data editing game, fieldworkers resolved inconsistencies best, data managers tied or ran a close second, and data users were the clear losers. We think this is reassuring given that field editing, in both its open and covert forms, is widespread in real-world fieldwork and presumably has been for a long time. In other words, the data that we use are probably better

after fieldworkers' editing than they would have been had many more unedited inconsistencies reached the ultimate data users.

More generally, however, our results suggest that at the very least, outsider researchers who remain concerned about field editing need to produce empirical evidence that their fears are warranted. The ball, to continue with our game metaphor, is in their court. Until they do, we will rely on our fieldworkers to resolve inconsistencies in the field. It appears to us to be at least as scientifically legitimate a choice as its outsider-oriented alternative.

Notes

1. "Data collection" is an established field of scientific inquiry that cuts across sociology, anthropology, psychology, and related disciplines (e.g., Hyman et al. 1954; Sudman and Bradburn 1974, 1982; Tanur 1992; Bernard 2000; Tourangeau, Rips, and Rasinski 2000; Maynard et al. 2002; Schaeffer and Presser 2003). "Data processing" is a scientific administrative task involving the transformation of raw data into usable data for research (Bourque and Clark 1992). Data editing may be embedded in either or both, as we show below.

2. Strong feelings, we believe, are common, as exemplified by two readers of an earlier version of this article. One of them, experienced in direct data collection, claimed that field editing is a "legitimate and established part of field operations." The other, by contrast, considered that computer-programmed editing is far superior to field editing and makes the latter unadvisable. More generally, we found field editing aversion to be most common among scientists with deep formal training in statistics and no field experience.

3. Throughout this article we limit ourselves to within-interview inconsistencies, not those that arise, for example, in differential husband–wife reports (e.g., Miller, Zulu, and Watkins 2001). The other perennial data problem that arises from data collection, missing values, is comparatively much less contentious since the availability of scientifically accepted techniques to deal with them statistically (Little and Rubin 1987, Allison 2001) makes it possible to let missing values remain missing for later treatment by the data user.

4. More specifically, all Current Population Survey (CPS) processing, including data cleaning, is contracted out to the Census Bureau by the Bureau of Labor Statistics, which requests edited CPS data (Greg Weyland, CPS survey statistician, personal communication, January 5, 2007). Consequently, CPS data appear to be consistent with the Census Bureau's definition of data quality as "fitness for use" (U.S. Bureau of the Census 2006).

5. The Demographic and Health Survey is a multiwave series of nationally representative surveys conducted in 78 countries since the late 1980s. Data and supporting information can be accessed at http://www.measuredhs.com.

6. These are both high-profile longitudinal research projects involving sociologists, demographers, economists, and anthropologists. For a description of the projects, visit http://lamp.opr.princeton.edu and http://malawi.pop.upenn.edu.

7. We use the term "callback" to refer to a second visit to a respondent who has already been interviewed, rather than, as it is otherwise used, to refer to repeated attempts to contact a respondent who has not yet been interviewed.

8. We not only assume that callbacks failed but also that probes embedded in the questionnaire—common in computer-assisted interviewing and widespread in U.S.-based surveys—did not solve the problem either. We address computer-assisted interviewing in the next subsection.

9. There is some variation among major users of computer-assisted interviewing with respect to how "hard" or "soft" the edits should be. The National Health Interview Survey, for example, includes both hard edits, which the interviewer must produce to continue the interview, and soft edits, which give the interviewer more discretion (Eve Powell-Griner, chief of the Data Analysis and Quality Assurance Branch, Division of Health Interview Statistics, National Center for Health Statistics, personal communication, January 10, 2007). In contrast, the CPS programs the survey instrument so as to allow the interviewer to undo edits and move on, even in the presence of inconsistencies, when errors cannot be fixed by probing (Greg Weyland, CPS survey statistician, personal communication, January 5, 2007).

10. The most publicized of these has been the so-called South Korean stem cell research scandal that exploded in the fall of 2005 and, among other consequences, led to the retraction of a breakthrough article in the top-tier peer-reviewed journal *Science* (Kolata 2005, *The Economist* 2006). Months later, a former University of Vermont professor became the first academic ever to receive a jail sentence in the United States for fabricating scientific data (Silverman 2006).

11. It would require substantial collaboration between the researcher and the directors and managers of one or more large survey projects. It would demand human resource and management skills to ensure that members of the field team do not feel uneasy with the presence of an outsider who appears to be evaluating their work. In addition, there is no guarantee that real data collection problems arising in a particular survey setting at a particular time will produce the type of sample structure that would facilitate the analysis. Finally, with the exception of cultural anthropologists, researchers themselves are rarely in the field for the duration of data collection.

12. We decided what information to alter in each selected questionnaire before seeing the questionnaires so that we had no influence on the selection process.

13. We are aware that this is a simplified, stylized scenario. As pointed out by a reviewer, many researchers participate in fieldwork and are attuned to the culture and context of the data collection setting. The same might also be said about researchers working in their own backyard—for example, American sociologists whose focus is middle-class Americans. That said, we believe that the detachment from the field of the data managers and data users in our experiment is a realistic representation of the real-world situation for most users of survey data, including the majority of first-world researchers whose analytic focus is developing countries. It also includes those of us who use, but have no hand in actually collecting, data on subpopulations within our own society with which we have no intimate familiarity.

14. All the information available for the restricted edits related to the household head and included basic demographic data, basic information on first and last domestic migration, first and last migratory trips to the United States, marital history, labor history, and additional data on the last migratory trip to the United States. In contrast, to make an unrestricted edit, the participants had access to the whole questionnaire and to all data from all other households in the community. The whole questionnaire included additional demographic data on the reference individual as well as domestic and international migration data for other members of the household; history of business, property, and land ownership for the household head; migration data on friends and relatives in the United States; history of illegal border crossings; ownership of vehicles, farm animals, and cattle; and house characteristics and amenities. In

practice, the participants rarely spent much time checking the additional data and virtually never looked at any data from other households.

References

Agar, Michael. 1980. *The Professional Stranger: An Informal Introduction to Ethnography*. New York: Academic Press.

Allison, Paul. 2001. *Missing Data* (Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-136). Thousand Oaks, CA: Sage.

Axinn, William G., Thomas E. Fricke, and Arland Thornton. 1991. "The Microdemographic Community-Study Approach: Improving Survey Data by Integrating the Ethnographic Method." *Sociological Methods & Research* 20:187-217.

Barber, Bernard. 1990. *Social Studies of Science*. Piscataway, NJ: Transaction Publishing.

Beatty, Paul. 1995. "Understanding the Standardized/Non-Standardized Interviewing Controversy." *Journal of Official Statistics* 11:147-60.

Bernard, H. Russell. 2000. *Social Research Methods: Qualitative and Quantitative Approaches*. Thousand Oaks, CA: Sage.

Blalock, Hubert M. 1961. *Causal Inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.

Bourque, Linda B. and Virginia A. Clark. 1992. *Processing Data: The Survey Example* (Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-085). Newbury Park, CA: Sage.

Cicourel, Aaron V. 1974. *Theory and Method in a Study of Argentine Fertility*. New York: John Wiley.

Collins, Randall. 2004. *Interaction Ritual Chains*. Princeton, NJ: Princeton University Press.

De Leeuw, Edith, Joop Hox, and Sabina Kef. 2003. "Computer-Assisted Self-Interviewing Tailored for Special Populations and Topics." *Field Methods* 15:223-51.

Dijkstra, Wil. 1987. "Interviewing Style and Respondent Behavior: An Experimental Study of the Survey Interviewer." *Sociological Methods & Research* 16:309-34.

"Egg on His Face: A Leading Stem-Cell Researcher Is Found Guilty of Fabricating His Results." 2006. The Economist, January 5, http://www.economist.com/science/displaystory.cfm?story_id=E1_VPVQSDS

Fletcher, Linda A., Darin J. Erickson, Traci L. Toomey, and Alexander C. Wagenaar. 2003. "Handheld Computers: A Feasible Alternative to Paper Forms for Field Data Collection." *Evaluation Review* 27 (2): 165-78.

Garfinkel, Harold. 1967. *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.

Gladwell, Malcolm. 2005. *Blink: The Power of Thinking Without Thinking*. Boston: Little, Brown.

Greene, Paul D. 2001. "Handheld Computers as Tools for Writing and Managing Field Data." *Field Methods* 13:181-97.

Haraway, Donna J. 1991. *Simians, Cyborgs, and Women: The Reinvention of Nature*. London: Routledge.

Heritage, John. 1984. *Garfinkel and Ethnomethodology*. Cambridge, UK: Polity Press.

Hyman, Herbert H., William J. Cobb, Jacob J. Feldman, Clyde W. Hart, and Charles Herbert Stember. 1954. *Interviewing in Social Research*. Chicago: University of Chicago Press.

Ice, Gillian H. 2004. "Technological Advances in Observational Data Collection: The Advantages and Limitations of Computer-Assisted Data Collection." *Field Methods* 16:352-75.

Jansen, Sue Curry. 1980. "The Stranger as Seer or Voyeur: A Dilemma of the Peep-Show Theory of Knowledge." *Qualitative Sociology* 2 (3): 22-55.

Kolata, Gina. 2005. "Amid Confusion, Journal Retracts Korean's Stem Cell Paper." *The New York Times*, December 31, http://www.nytimes.com/2005/12/31/science/31clone.html

Leahey, Erin. 2004. "The Role of Status in Evaluating Research: The Case of Data Editing." *Social Science Research* 33:521-37.

———. Forthcoming. "Overseeing Research Practice: The Case of Data Editing." *Science, Technology & Human Values*.

Leahey, Erin, Barbara Entwisle, and Peter Einaudi. 2003. "Diversity in Everyday Research Practice: The Case of Data Editing." *Sociological Methods & Research* 31:63-89.

Little, Roderick J. A. and Donald B. Rubin. 1987. *Statistical Analysis With Missing Data*. New York: John Wiley.

Martinson, Brian C., Melissa S. Anderson, and Raymond de Vries. 2005. "Scientists Behaving Badly." *Nature* 435:737-8.

Massey, Douglas S. 1987. "The Ethnosurvey in Theory and Practice." *International Migration Review* 21:1498-522.

Maynard, Douglas W., Hanneke Houtkoop-Steenstra, Nora Cate Schaeffer, and Johannes van der Zouwen, eds. 2002. *Standardization and Tacit Knowledge: Interaction and Practice in the Survey Interview*. New York: John Wiley.

Mensch, Barbara S., Paul C. Hewett, and Annabel Erulkar. 2003. "The Reporting of Sensitive Behavior by Adolescents: A Methodological Experiment in Kenya." *Demography* 40:247-68.

Miller, Kate, Eliya M. Zulu, and Susan C. Watkins. 2001. "Husband–Wife Survey Responses in Malawi." *Studies in Family Planning*, 32 (2): 161-74.

Nash, Dennison. 1963. "The Ethnologist as Stranger." *Southwestern Journal of Anthropology* 19:149-67.

ORC Macro. 2002. *Supervisor's and Editor's Manual for Use With Model A nd B Questionnaires* (Measure DHS + Basic Documentation No. 5). Calverton, MD: Author.

Powdermaker, Hortense. 1966. *Stranger and Friend: The Way of the Anthropologist*. New York: W. W. Norton.

Prewitt, Kenneth. 1983. "Management of Survey Organizations." Pp. 123-44 in *Handbook of Survey Research*, edited by Peter H. Rossi, James D. Wright, and Andy B. Anderson. New York: Academic Press.

Rubenstein, Ariel. 2006. "Freak-Freakonomics." *Economists' Voice* 3 (9): Article 7. http://www.bepress.com/ev/vol3/iss9/art7

Schaeffer, Nora C. and Stanley Presser. 2003. "The Science of Asking Questions." *Annual Review of Sociology* 29:65-88.

Schober, Michael and Frederick G. Conrad. 1997. "Does Conversational Interviewing Reduce Survey Measurement Error." *Public Opinion Quarterly* 61:576-602.

Schwartz, Barry. 2004. *The Paradox of Choice: Why More Is Less*. New York: HarperCollins.

Silverman, Adam. 2006. "Former UVM Professor Sentenced to Jail for Fraud." *Burlington Free Press*, June 29, http://www.burlingtonfreepress.com/apps/pbcs.dll/article?AID = 20060629/NEWS01/606290315/1009

Stocking, George W. 1983. "History of Anthropology: Whence/Whither." Pp. 3-12 in *Observers Observed: Essays on Ethnographic Fieldwork*, edited by George W. Stocking. Madison: University of Wisconsin Press.

Suchman, Lucy and Brigitte Jordan. 1990. "Interactional Troubles in Face-to-Face Survey Interviews." *Journal of American Statistical Association* 85:232-53.

Suchman, Mark C. 1995. "Managing Legitimacy: Strategic and Institutional Approaches." *Academy of Management Review* 20:571-610.

Sudman, Seymour and Norman M. Bradburn. 1974. *Response Effects in Surveys: A Review and Synthesis*. Chicago: Aldine.

———. 1982. *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco: Jossey-Bass.

Tanur, Judith M., ed. 1992. *Questions About Questions: Inquiries Into the Cognitive Bases of Surveys*. New York: Russell Sage Foundation.

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The Psychology of Survey Response*. Cambridge, UK: Cambridge University Press.

U.S. Bureau of the Census. 2002. *Current Population Survey: Design and Methodology* (Technical Paper 63RV). Washington, DC: Author.

———. 2006. *Definition of Data Quality. Version 1.2*. Washington, DC: Census Bureau Methodology & Standards Council.

Van den Broeck, Jan, Solveig Argeseanu Cunningham, Roger Eeckels, and Kobus Herbst. 2005. "Data Cleaning: Detecting, Diagnosing, and Editing Data Abnormalities." *PLoS Med* 2 (10): e267.

Weber, Max. [1922] 1968. *Economy and Society*. New York: Bedminster.

Weinreb, Alexander A. 2006. "The Limitations of Stranger-Interviewers in Rural Kenya." *American Sociological Review* 71:1014-39.

Zeller, Richard A. and Edward G. Carmines. 1980. *Measurement in the Social Sciences: The Link Between Theory and Data*. New York: Cambridge University Press.

Mariano Sana is an assistant professor of sociology at Louisiana State University with years of experience overseeing field operations for a National Institutes of Health–funded research project in various Latin American countries. Most of his research focuses on international migration. Most recently, he has published in *Social Forces*, *Social Science Quarterly*, and *Population Research and Policy Review*.

Alexander A. Weinreb is a lecturer in the Department of Sociology and Anthropology at Hebrew University. He is a social demographer who, for a number of years, was field director of National Institutes of Health–funded longitudinal surveys in Kenya and Malawi. His recent publications have appeared in the *American Sociological Review*, *Population and Development Review*, and other leading journals.