Surveys are the main source of information about poverty, health, demography and many other indicators in Africa, making them vital for evidence-based policy design and planning. But have social scientists been ignoring the potential impact of interviewers on the data they collect? A new study[1], whose researchers include LSE’s Ernestina Coast, highlights just how much influence interviewers have over the data they produce.

Many social scientists analyse household survey data collected by others. Though this early stage of the research process is often out of our hands, we need to consider how the data were produced as it may have serious consequences for their interpretation. Our research identifies situations in the production of household survey data where the interviewer has considerable influence over the data collected.

Censuses and household surveys in Africa remain almost totally dependent on interviewers for data collection. Data quality is ensured by good quality training and supervision, along with comprehensive interviewers’ manuals. Nevertheless, the interviewer has significant power to negotiate access to respondents and influence the responses they provide. Ultimately, this matters because it affects the validity of the data; data that are then used to allocate scarce resources.

The improved availability of standardised international data sets means that increasing numbers of researchers with no fieldwork experience are working on secondary data analysis. The concern is that they may be accepting these data at face value, paying little attention to any difficulties encountered in the field. Our paper gives the specific stages in the data collection process where data users may wish to reflect on how this could influence their analysis and conclusions. We identify key areas in which concept ambiguities may have a strong influence on data validity.

This kind of reflection is crucial for interpreting data, but may be easily ignored, potentially resulting in misleading conclusions. Take, for example, the responses to questions on “ideal family size”. Data on fertility are used by policy designers and planners interested in how many children women actually want, whether they are able to achieve this and how this is changing over time. The main source of information on fertility intentions comes from the Demographic and Health Surveys. While women might reply “It is up to God” when asked how many children they would choose to have, instead interviewers coax the woman into giving a numeric response. This is where the power of the interviewer is exerted. We found considerable room for interpretation and manoeuvre, as interviewers push the respondent and influence the outcome, frequently in a way that demonstrates their own opinions about what counts as a suitable reply. One woman, persuaded by an interviewer into giving a numeric response to this question suggested 20 children. The interviewer then queried this asking how she would afford to feed them. The woman lowered her number to 8. How meaningful is this response? Where the respondent is ambivalent or uncertain he or she is likely to be induced to provide an answer which the interviewer finds acceptable. This influence over responses was just one of the ways in which we found interviewers to have direct power over the data produced.

This study is the first of its kind, following statistical data all the way along the chain of production right through to how it is used by government planners and analysts. We conducted the mixed methods study in four sub-Saharan
countries: Tanzania, Senegal, Burkina Faso and Uganda. We interviewed key informants along the chain of data production and exploitation: interviewers, supervisors, analysts, statisticians, academics, survey designers, commissioners and data users. We reviewed methodological materials for recent censuses and national surveys in all four countries. Finally, we interviewed households in each country to explore respondents’ perceptions of their own household / domestic group membership, and compared how well this matched the definitions found in the survey manuals. This paper is based on work from two research projects[2] that investigated the definition of the household, and is complemented by insights from long term demographic quantitative and qualitative research by the authors in sub-Saharan Africa.

While interviewers we spoke to were often articulate about their power and the decisions they make, data users, however, who had never worked as interviewers, admitted that they did not question how the data were produced and the negotiations necessary to produce them. Analysts’ ignorance both of the actual parameters of definitions and of the different strategies and skills required to obtain their data may well lead to unreliable analyses and conclusions. We must not lose sight of the immense power of the interviewer over the quality of the raw data, and thus the analyses and any subsequent policy decisions.

Read the full paper


Related