USAID
FROM THE AMERICAN PEOPLE

# Data Quality in Demographic and Health Surveys That Used Long and Short Questionnaires (MR30)

*An Analysis Brief from The DHS Program*

## Why study DHS questionnaire length?

Since the first phase of The DHS Program in 1984, DHS surveys have increased in scope and complexity as questionnaires are lengthened and survey modules are added. Increase in questionnaire length can increase interview time, and ultimately may present more burden for both the interviewer and the respondent.

Many DHS surveys have included modules or additional topic-specific sections only among subsamples, resulting in questionnaires of differing lengths within the same survey. It seems intuitive that longer questionnaires would have different effects than shorter questionnaires on fieldwork, interviewer fatigue and performance, and survey implementation. Surveys that have two different lengths of questionnaires offer an opportunity to explore the extent to which questionnaires of different lengths may have these effects.

## Which countries were included in the study?

This analysis included data from the 2016 South Africa DHS, 2014 Kenya DHS, and 2015-16 India National Family Health Survey (NFHS), as these surveys were designed with two versions of the same questionnaire – one version considerably longer than the other. Additionally, key informant interviews with survey experts who worked on surveys with long and short questionnaires were conducted to understand how these surveys are implemented, and data quality considerations, if any, along the survey process.


## What methods were used to conduct this analysis?

This report used mixed methods to investigate data quality and implementation differences in long and short questionnaires. To understand the effect, if any, of questionnaire length on data quality, two types of data quality indicators were examined: indicators that may reflect efforts on the part of fieldworkers to reduce survey burden (i.e., their workload) and those concerning age and date of birth that are notoriously difficult to collect accurately in household surveys, though not all indicators were available for each survey.

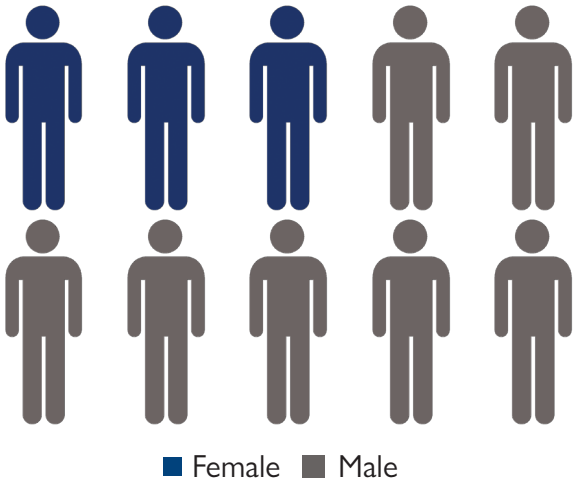
- **Age displacement ratios:** Interviewers may be tempted to reduce the number of women eligible for the individual interview in order to reduce their overall workload by shifting women's ages outside the standard DHS range of age 15-49. The ratio of female household members age 14 to female household members age 15 checks the lower boundary of age eligibility for age displacement, while the ratio of female household

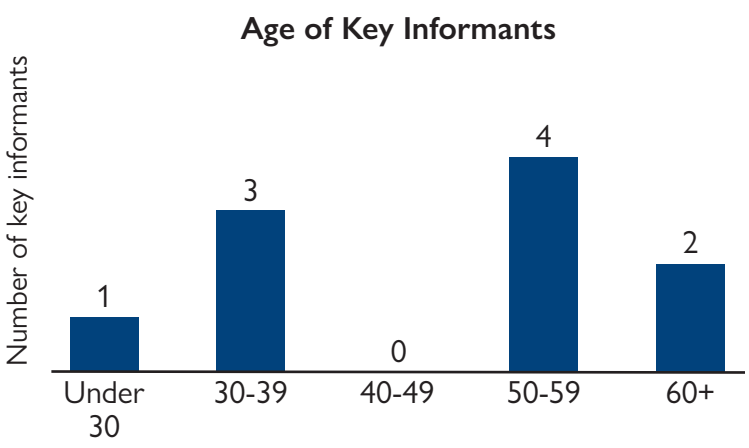
members age 50 to female household members age 49 checks the upper boundary.

- **Average number of eligible persons in a household:** The average number of women eligible for the individual interview in each household and the average number of eligible children under age 5 for height and weight measurement in each household can also reveal where efforts may have been taken to reduce the number of eligible persons.
- **Skipping questionnaire sections:** Filter questions determine whether an entire section or series of questions are skipped or asked of the respondent. For example, if a woman says that she has never heard of HIV, she is not asked any of the HIV/AIDS-related knowledge, attitude, and behavior questions. In addition to knowledge of HIV, filter questions on use of a family planning method and sufficient privacy to conduct the domestic violence module were examined.
- **Completeness of date of birth:** Interviewers are expected to make serious efforts to help respondents recall at least the month and year for dates of birth. There may be temptation, when interviewers know they are facing a longer interview, to conserve energy, spend less time probing, and accept more missing data. Completeness of date of birth was examined for births in the five years preceeding the survey and for women.

**Key Informant Interviews:** Additionally, this study explores themes resulting from qualitative interviews with 10 survey experts who worked on surveys with long and short questionnaires to understand:

- How are long and short questionnaires implemented in a survey and are there other elements, advantages, or disadvantages to consider when administering long and short questionnaires in the same survey? How is data quality a consideration along the survey process?




## What are the key results?

Major results from the quantitative anaylsis include:

- The long questionnaires in each country had large differences in the average number of variables per woman compared to a country's short questionnaires. Despite these differences, there is little evidence that interviewers made different efforts to reduce their workloads based on whether they were administering the long or the short questionnaire.

Major qualitative themes from the key informant interviews include:

## The case for having two questionnaires

Having a long and short questionnaire makes it possible to meet survey stakeholders' needs for specific data while keeping survey cost within the available funding. It can also provide an opportunity to improve on past choices.

**"So a lot of times outside voices, other funders, other public health players at the table, want things added to a survey for their own use or their own programmatic needs that may… be valid interests, but really do add to the burden of a survey... If you do a long and a short version, then you can still receive the funding that they bring, meet their data needs, but lessen the burden on your field staff and on other quality and logistical concerns. So I think it allows more flexibility while also maybe being able to still fund a survey through these donors, who have demands."**

## Decisions about survey and questionnaire design

Decisions about the sampling for long and short questionnaires are straightforward, but decisions about what content to include in which questionnaire can be complicated, and there will be extra considerations when time comes to finalize the data.

**"…it's more about what you would like to measure. What are the key indicators, from these two surveys, from these two questionnaires and what is your target population, or what is the denominator of your indicator? And so, these are the two main questions that guide our decisions when it comes to the distribution, or I'll call it the allocation of the sampling units, over the sampling strata."**

## Survey burden

Having a long and short questionnaire results in a survey that is easier to manage; both fieldworker training and fieldwork monitoring are largely unaffected. However, some consideration is required during the day-to-day implementation of fieldwork (which questionnaires are tackled first, which questionnaires are assigned to which interviewers), and long questionnaires are more fatiguing to interviewers and respondents.

**"I don't think I noticed any [differences] necessarily in implementation or in energy level, except for that, when [interviewers] knew they had a long women's questionnaires, they would have really hunkered down… And when they had a short questionnaire, it was just like – I got this, don't worry."**

## Decisions about survey and questionnaire design

There is no set definition for what constitutes a long and short questionnaire. If survey stakeholders do not fully understand the effect of the different sampling in long and short questionnaire surveys on indicator calculations, they may be left unsatisfied with the survey results.

**"… when people then try and analyze data… until they get really familiar and read all the [documentation] they don't realize that certain indicators are not for the [lower administrative unit] level."**

### Data quality

Data quality and its relationship with survey processes is an important consideration throughout the design and implementation of surveys with long and short questionnaires. There exists a conventional wisdom of an inverse relationship between questionnaire length, among other survey elements, and data quality, and certain data quality indicators can be monotored to identify issues during fieldwork.

> **"Or if the questionnaires are too heavy, you try to reduce the content. This is also [a] data quality concern. You reduce the workload of the interviewers; you expect that to improve or to get a better data quality."**

### Opinion and Future

While some key informants had reservations about having long and short questionnaires during survey design, once the implementation began, they observed their value and their usefulness for future surveys, either in setting a precedent or to find solutions for future survey design.

> **"This is an interesting solution to the challenge of everyone wanting all the data."**

## Conclusions and Recommendations

This analysis found little evidence that having differing lengths of questionnaires resulted in data quality differences between the resulting data from the two questionnaires in the surveys examined. Key informants agree that deploying long and short questionnaires solves a significant problem in survey design and implementation and were unanimously supportive of using long and short questionnaires in the future. For future surveys, particularly those where it is desirable to provide at least some indicators for lower administrative levels or nonstandard populations, the authors recommend the following:

- The long and short questionnaire approach should be included among survey design options
- Tabulations should be developed at the beginning of the survey process for long and short questionnaires (or other country-specific subsampling) and shared with stakeholders
- Fieldwork should be structured around the length of questionnaires
- Long and short questionnaires should be equally distributed among interviewers and interviewers should be coached to implement long questionnaires with confidence

This brief summarizes The DHS Program's Methodological Report 30, by Courtney K. Allen, Julia Fleuret, and Jehan Ahmed with funding from The United States Agency for International Development through The DHS Program implemented by ICF. The full report is available at https://www.dhsprogram.com/publications/publication-mr30-methodological-reports.cfm