

Module 6: Data quality monitoring

Video 1 of 6: Module Overview and Introduction

Sharan Sharma



THE WORLD BANK

MANNHEIM
BUSINESS SCHOOL

November 2020

Module 6 – Remote training on Phone Surveys

At the End of the Module Participants Should...

- Be familiar with the various data sources used for quality monitoring.
- Be able to assess the relative strengths and weaknesses of each data source used for quality control.

Module design

- Introduction
- Overview of Data, Tools, and Methods
 - Paradata and Computer audio-recorded interviewing (CARI)
- Practical quality control

Why bother monitoring data quality?

1. Were our assumptions/design decisions correct?

- Especially true for new surveys or surveys which have undergone changes, e.g., due to COVID-19

2. Are we getting what we set out to measure?

- Total Survey Quality (Total Survey Error → Module 1)
- Quality assurance-Quality control
- Ethical and Legal issues (e.g., preventing falsification)

...Why bother monitoring data quality?

3. Should we make a change to our design or implementation?

- Adaptive designs
- Do we need to retrain interviewers?

4. What should we change in our next round of data collection, if anything?

- Continuous improvement
- Do we need to reallocate field resources?

...Why bother monitoring data quality?

5. Fieldwork is contracted-out for many international surveys

Lesser control on interviewer selection and training → more effort on monitoring.

6. Interviewers do their job better when they know they are being monitored + value their training more when they know they are being monitored (*Fowler and Mangione, 1990*)

What exactly should we monitor?

- Sampling errors
 - Non-sampling errors
 - Specification errors
 - Coverage errors
 - **Non-response errors**
 - **Measurement errors**
 - Processing errors
- } Focus of this module

...and these errors can stem from multiple sources...

- Questionnaire
 - Respondent
 - **Interviewer** ← Focus of this module
 - Geography
- ...etc.

Unit non-response

Interviewer 1

Interviewer: Could I please speak to Mr. Shah?

Respondent: Yes, this is Mr. Shah.

Interviewer: Sir, my name is Dinesh and I am calling on behalf of ABC survey company. May I speak to you for a brief survey?

.....

Interviewer 2

Interviewer: Could I please speak to Mr. Shah?

Respondent: Yes, this is Mr. Shah.

Interviewer: Sir, my name is Dinesh and I am calling on behalf of ABC survey company. May I speak to you for a brief survey?

.....

Interviewer 1

....

Respondent (in an irritated tone): I really cannot do this.

Interviewer: Sorry to hear that, sir. Have a nice day.

Interviewer 2

...

Respondent (in an irritated tone): I really cannot do this.

Interviewer: Sorry to hear that sir. It will just take 15 minutes.

Respondent : I am really busy! You called me in the middle of some work !

Interviewer: I sincerely apologize for this. It would be great if I could call you sometime else as per your convenience. You have been specially chosen to answer this survey which is of national importance.... your response matters a lot ...

Interviewer 1

....

Interviewer 2

...

Respondent: Okay, try me over the weekend.

Interviewer: Thank you very much, Mr. Shah. I'll call you on Sunday at 11 am. I hope that would be fine?

Respondent: Ok.

Interviewer: Great. Thank you. And once again, sorry for disturbing you.

Item non-response

Interviewer 1

Interviewer: What was your gross salary in the year 2019 ?

Respondent: I don't want to disclose this.

Interviewer: ok, moving into the next question,...

Interviewer 2

Interviewer: What was your gross salary in the year 2019 ?

Respondent: I don't want to disclose this.

Interviewer: I understand your concern. As I said in the beginning, your responses will be completely confidential. All data are going to be used only in aggregate – your responses are not going to be singled out.

Why monitor non-response?

- We never know what the non-respondents would have actually said.
- If responses (from respondents) are systematically different from answers that non-respondents would have given, then we may end up with biased estimates.

What exactly should we monitor?

- Sampling errors ?
- Non-sampling errors ?
 - Specification errors ?
 - Coverage errors
 - Non-response errors ?
 - **Measurement errors** ?
 - Processing errors ?

Perhaps the most
damaging ?



END OF VIDEO 1