
Photo: Selam Desta, PMA2020

PMA2020

Performance Monitoring and Accountability 2020

Johns Hopkins Bloomberg School of Public Health | Bill & Melinda Gates Institute for Population and Reproductive Health

Questionnaire answer agreement: Detect lurking problems and increase data quality

James Pringle, ScM
Sr. Programmer Analyst
Bill & Melinda Gates Institute for Population and Reproductive Health

CSDI Workshop, 28 March 2018

Our Vision:

To meet the evolving health and development data needs of country stakeholders across low and middle income countries through innovative approaches to rapid data collection.

Our Mission:

We revolutionize the way global family planning and health data is collected and promote its effective utilization. We do this by building capacity of country-level partners to design and implement innovative, high-quality data-collection platforms that empower stakeholders to make informed policy and program decisions.

PMA2020

Photo: Dr. Linnea Zimmerman, PMA2020

PMA2020 COUNTRIES/PARTNERS:

The problem

How does a data analyst know that what is in the dataset is the truth?

There are several sources of ***measurement error*** in mobile data collection on the ***interviewer side***. Examples are:

- Being unfamiliar with the software
- Similar physical movements between tapping a button to record an answer and navigating to the next question
- Factors unrelated to technology such as not probing to clarify an answer, mishearing, or difficulty classifying a response

The experiment

- In order to investigate the problem of measurement error on the interviewer side, we analyzed surveys where multiple interviewers were present and separately recorded answers.
- Ideally, all recorded answers would be the same.
- There could be differences. Considering each recorded answer as a vote, the answer with the most votes is most likely to be true.

Definitions

- Within a group, a single question has ***agreement*** if all group members answer the same way. A question that does not have agreement has ***disagreement***.
- The ***total agreement*** for a group is the percent of all questions that have ***agreement***.

Gathering the data

PMA2020 trainings provided the ideal setting for the experiment since supervisors and enumerators were learning new content and practicing in the classroom and in the field. The data comes from:

- 10/2017 Ghana Primary Health Care module training, facility questionnaire
- 02/2018 Nigeria Abortion module training, facility questionnaire
- 03/2018 India Adolescent Health and Abortion modules training, female and facility questionnaire

Ghana experiment

- Fifteen supervisors broke into three teams of five.
- The experiment was explained to them.
- Each team visited a different health facility and administered the questionnaire together to a single respondent.

Ghana experiment (cont'd)

Team	Total agreement
1	100%
2	100%
3	100%

Ghana experiment (cont'd)

Team	Total agreement
1	87%
2	84%
3	89%

Ghana experiment (cont'd)

Questions with at least one team disagreeing

- 50 answers had at least one team disagreeing
- 166 total answers (points of comparison between the teams)
- 50/166 (30%) of all answers had at least one team disagreeing

Ghana experiment (cont'd)

Questions with all teams disagreeing

PHC 111. May I see a nearby handwashing facility that is used by staff? *Observe*

- Soap is present
- Stored water is present
- Running water is present
- Handwashing area is near a sanitation facility
- None of the above
- Did not see the facility.

PHC 904. Does your facility maintain records to track revenue and expenditures?

- Not Available
- Reported, but unseen
- Observed
- Do not know
- No response

Nigeria experiment

- In Nigeria, the supervisors for the abortion pilot broke into groups to visit health facilities and ask about abortion services.
- I only have survey submissions from two groups. One for a hospital and one for a pharmacy.

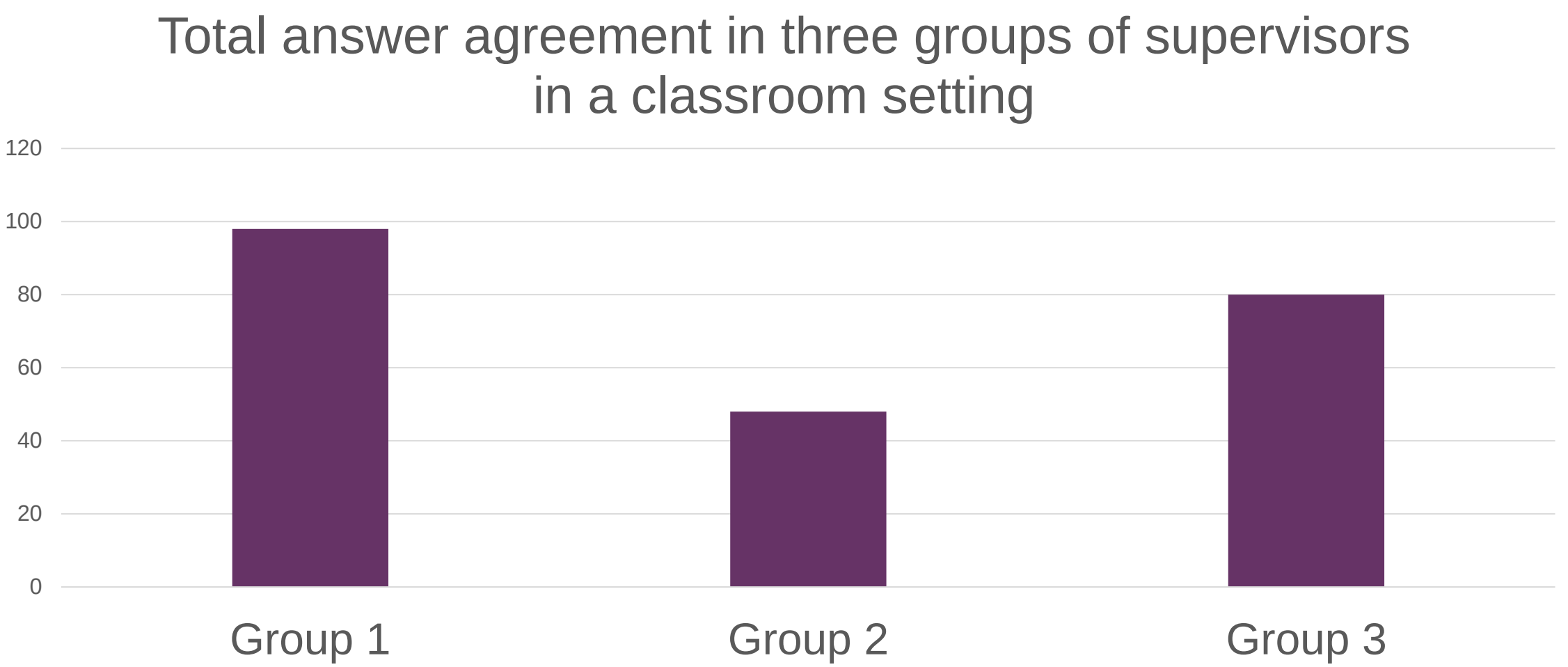
Team	Total agreement
Hospital	95%
Pharmacy	100%

India experiment

- Three supervisors went to a health facility to administer approximately thirty abortion-related questions.
- Ten female enumerators visited women and adolescents in the surrounding village to administer adolescent health and abortion-related questions.
- Of the female enumerators, only one group of three followed instructions, so we have data from their four interviews.

Team	Total Agreement
Health center	88%
Female survey #1	90%
Female survey #2	95%
Female survey #3	94%
Female survey #4	98%

India experiment



India experiment (cont'd)

Question with disagreement

- SQ 609. Which of the following methods are used in this facility for post-abortion care or abortion for pregnancies of 12 weeks or less?
  - Misoprostol alone
  - Mifepristone and misoprostol alone
  - Manual vacuum aspiration (MVA)
  - Electric vacuum aspiration (EVA)
  - Dilation and evacuation (D&E)
  - Dilation and curettage (D&C)
  - Laparotomy

India experiment (cont'd)

- Different methods of abortion, but similar.
- ***D&C:*** A procedure where the cervix is dilated and a special instrument is used to scrape the uterine lining
- ***D&E:*** A procedure where the cervix is dilated and fetal tissue is removed using surgical instruments (such as forceps and curette) and suction

Try this out yourself

The analysis code used to generate numbers and figures for this presentation is available online. This code can produce new analyses as well. The link is:

github.com/jkpr/AnswerAgreement

What is needed:

(1) A survey dataset with rows as individual survey responses and columns as headers to different questions, (2) Optionally, a column identifying groups, (3) Optionally, a list of columns for which you want to check agreement

Limitations

- While groups of two interviewers could work, you need three to have a tiebreaker if someone is off.
- It was difficult to get all the data after it was collected since submitting data normally requires Internet access for Interviewers.
- It was hard for me to get into the schedule to run this activity. We do a lot of important things at our trainings!
- Analysis code is written in Python. It's free, but difficult to teach others a new tool.

Next steps

- Continue to build out the features of the analysis code. Make it easier to get the results that are interesting.
- Take majority rule to be the “truth” and calculate individual accuracies within groups.

Conclusions

- It is easy and natural to incorporate answer agreement experiments into trainings leading up to data collection
- Providing quick analysis and feedback is an opportunity to detect content matter that may need clarification
- The answer agreement experiment itself is a strong lesson on being deliberate and careful when recording answers
- Measurement error likely exists in all survey datasets, and we must not lose skepticism during analysis

Contact me

James Pringle

jpringle@jhu.edu

github.com/jkpr

Performance Monitoring and Accountability 2020

Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe Street, Baltimore, MD 21205

www.pma2020.org

facebook.com/GATES.PMA2020 Twitter: @PMA2020JHU