**INSTITUTE FOR SOCIAL RESEARCH**
**SURVEY RESEARCH CENTER**
UNIVERSITY OF MICHIGAN

# CAPI, Web-based Survey Management, and Interviewing Quality Control: Illustrations from Nepal (and elsewhere)

William G. Axinn
Stephanie Chardoul

IUSSP
December 5, 2019

# Overview

- Computer-assisted interviewing (CAI) helps to reduce interviewer error, and provides paradata for better management of surveys

- These tools reduce <u>error</u> – increasing efficiency – and they reduce <u>bias</u> – increasing accuracy

- U-M's Survey Research Center collaborates with global partners to apply best practices in survey design, including CAI technical systems

  Recent: China, Ghana, India, Nepal, Saudi Arabia, Qatar

# Review of TSE

- Total Survey Error is the sum of all ways survey measurement can go wrong
  - Differences between true and measured error
- Variance, or variable, error: random; no expected impact on mean values but reduced efficiency
- Bias, or systematic, error: directional and alters mean estimates (and model estimates)
- CAI systems can be designed to capture and reduce interviewer-originated error

# International Surveys

- Pressure to compete for higher quality and lower cost
- Previous presentation focused on sample bias and adaptive design; we will focus on **measurement quality**
- Quality control and quality assurance programs still not well developed in international surveys
  - Organizations lack financial, methodological, & technical resources and expertise
- QC/QA applied at three levels
  - Survey product
  - Survey process
  - Survey organization
- Need to develop accessible tools for each

# Crucial for Success of Population-Scale Measurement

- Reduction of bias is essential to the success of understanding key questions
  - Specific population problems could go undetected
  - Measurement of sub-group differences can be in error
  - Measures of change can over-estimate program effect
  - Predictive models can yield false conclusions
- Even reduction of random error supports detection of differences, trends, and changes produced by policies or programs
  - Greater efficiency means fixed fieldwork budgets produce more reliable measurement

# Benefits of CAI (1)

- Preload of sample frame or prior wave data
  - Confirmation of correct respondent; dependent interviewing (confirmation of known information)
- Correct implementation of sample rules
  - Application of eligibility criteria
  - Generation of "spawned" new sample lines
- Explicit interviewer instructions/checkpoints
- On-line access to question-level help files
- Programmed item-level range and consistency checks

# Benefits of CAI (2)

- Correct display of text fills in question text

- Programmed survey path logic/skip patterns

- Real-time generation of constructed variables

- Systematic collection of interviewer observations

- Systematic collection of respondent contact (and other) information

- Generation of paradata from both sample management and questionnaire systems

# CAI "Enabled" Quality Control (and Production Management)

- Assumes electronic sample management system, as well as computerized questionnaire

- Controlled assignment of sample to interviewers

- Real time (almost) access to contact histories/status – including time stamps

- GPS confirmations

- Questionnaire/survey time stamps (overall, section-, and item-level)
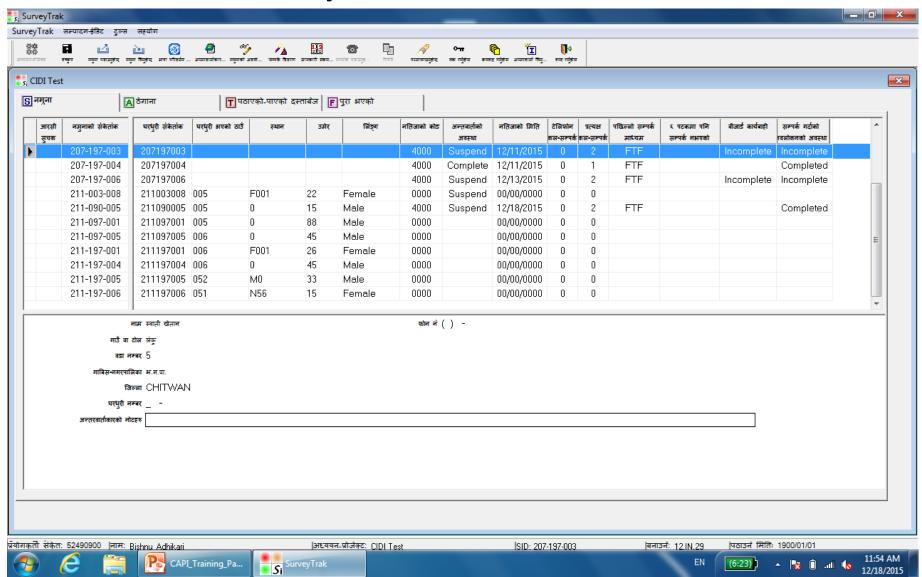
- Questionnaire/survey data

# Case Study: Nepal

- Chitwan Valley Family Study
  - Longitudinal since 1997
  - 10,000+ individuals
  - Continuous Household Registry
  - Periodic household and individual data collections
- Transitioned Household Registry from PAPI to CAPI in 2015
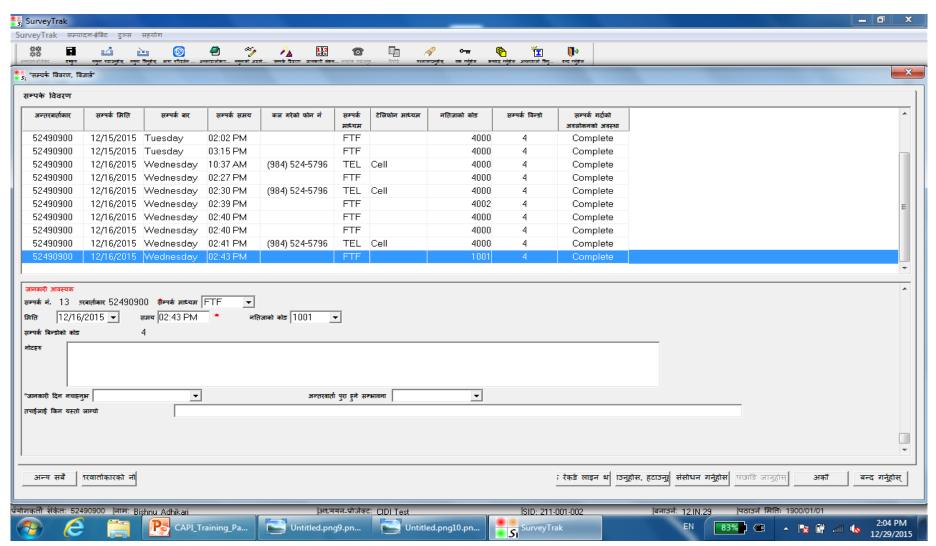- Used CAPI for complex mental health + genetics data collection 2016-2018

# SurveyTrak International

# SurveyTrak International

# CAI Questionnaire



प्रश्न खण्ड

बैकल्पिक उत्तर खण्ड

उत्तर खण्ड

12

# Nepal

- Interviewers used SurveyTrak to record outcome of **every** contact attempt

- Transmitted via internet daily to U.S.; information across interviewers compiled into master reporting dataset

- Web-based management tool ("WebTrak") used to provide reports to production management team in Nepal

- SRC team taught Nepal team how to review reports and identify potential problems

**Table 1: Data Collection Status**

| S.N. | Status | Number | Percentage | Remarks |
|------|--------|--------|------------|---------|
| 1 | Interview Completed | 2549 | 89.63 | |
| 2 | Interview Incomplete due to different reasons (Tracking) | 295 | 10.15 | |
| 3 | Interview Incomplete but Finalized (No more eligible for interview) | 63 | 2.17 | |
| | **Total Visited Resp** | **2907** | **100** | |

**Table 2: Interview Length by Data Collection Mode**

| S.N | Mode | Minimum | Maximum | Average |
|-----|------|---------|---------|---------|
| 1 | Face to Face | 32.68 | 398.2 | 79.27 |
| 2 | Telephone | 42.46 | 247.7 | 74.24 |

**Table 3: Interview Length by Resp. Gender**

| S.N | Gender | Minimum | Maximum | Average |
|-----|--------|---------|---------|---------|
| 1 | Female | 37.86 | 398.2 | 85.06 |
| 2 | Male | 32.7 | 247.7 | 70.14 |

**Table 4: Interview Length by Resp Age group**

| S.N | Age Group | Minimum | Maximum | Average |
|-----|-----------|---------|---------|---------|
| 1 | <18 | 39.7 | 192.9 | 66.0 |
| 2 | 18-34 | 32.7 | 398.2 | 74.9 |
| 3 | 35-49 | 43.4 | 272.4 | 84.2 |
| 4 | 50-59 | 44.7 | 235.2 | 89.2 |

**Table 5: Saliva Collection and logged in central office**

| S.N | Status | Number | Percentage | Remarks |
|-----|--------|--------|------------|---------|
| 1 | Yes | 1260 | 64.2 | |
| 2 | No (Refused) | 5 | 0.3 | |
| 3 | Saliva yet to be collected (No | 699 | 35.6 | |
| 4 | Logged in central office | 1260 | 100.0 | |
| | **Total** | **1964** | **100** | |

**Table 1: Timing Summary by interviewer (from Merged Data)**

| S.N | Interviewer | Interviewer ID | Total IW | IW Time (Minutes) | | | % IW |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | | | | Min | Max | Mean | |
| 1 | | 52410702 | 57 | 43.9 | 101.8 | 63.7 | 2.2 |
| 2 | | 52410704 | 49 | 50.4 | 131.3 | 74.4 | 1.9 |
| 3 | | 52410710 | 13 | 39.7 | 131.3 | 72.2 | 0.5 |
| 4 | | 52410717 | 127 | 40.9 | 172.4 | 70.7 | 5.0 |
| 5 | | 52410718 | 141 | 44.0 | 220.8 | 91.6 | 5.5 |
| 6 | | 52410719 | 20 | 60.1 | 193.8 | 82.4 | 0.8 |
| 7 | | 52410723 | 34 | 57.4 | 149.5 | 75.2 | 1.3 |
| 8 | | 52410726 | 153 | 42.4 | 150.5 | 65.2 | 6.0 |
| 9 | | 52410727 | 159 | 46.3 | 171.0 | 70.6 | 6.2 |
| 10 | | 52410728 | 178 | 44.8 | 139.6 | 73.9 | 7.0 |
| 11 | | 52410729 | 165 | 45.2 | 144.6 | 72.0 | 6.5 |
| 12 | | 52410730 | 149 | 39.7 | 234.5 | 59.7 | 5.9 |
| 13 | | 52410731 | 78 | 45.6 | 168.8 | 75.1 | 3.1 |
| 14 | | 52410734 | 172 | 47.5 | 264.8 | 88.8 | 6.8 |
| 15 | | 52410735 | 70 | 45.1 | 127.3 | 71.0 | 2.7 |
| 16 | | 52410736 | 151 | 42.6 | 229.5 | 93.4 | 5.9 |
| 17 | | 52410737 | 168 | 43.5 | 158.3 | 78.4 | 6.6 |
| 18 | | 52410740 | 147 | 48.2 | 199.0 | 68.1 | 5.8 |
| 19 | | 52410746 | 28 | 60.9 | 178.1 | 97.8 | 1.1 |
| 20 | | 52410747 | 179 | 37.8 | 236.8 | 93.3 | 7.0 |
| 21 | | 52410748 | 169 | 45.9 | 141.6 | 73.2 | 6.6 |
| 22 | | 52410750 | 139 | 47.3 | 390.4 | 92.3 | 5.5 |
| | **Total** | | **2546** | **37.85** | **390.38** | **77.56** | **100.0** |

**Table 2: Number of questions asked per IW, Average time, questions asked less than a second by Interviewer (ADT Analysis)**

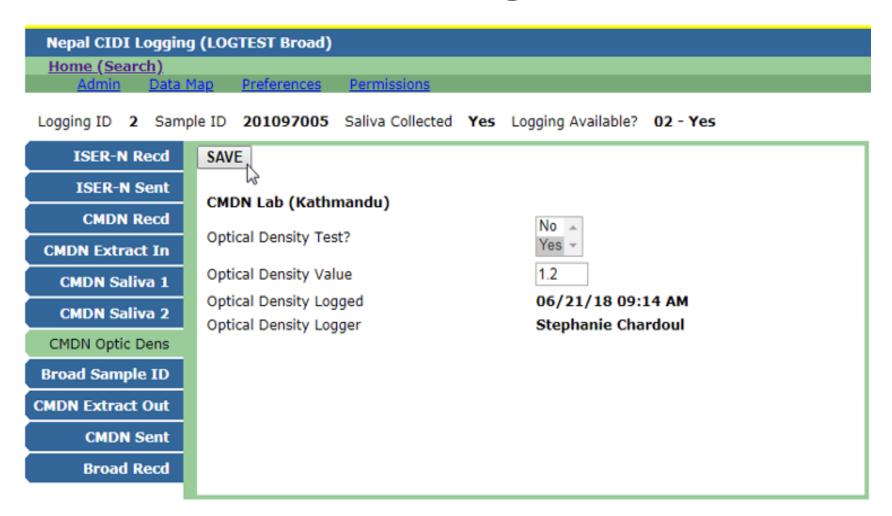| S.N | Iwer's ID | Comp_IW | Avg. Call | Average of SessionDuration_minutes | Avg Que. Asked | Avg. DK Resp/IW | Avg. RF Resp/IW | Avg. of Question <1 sec | LongPause GT10 min | F2Count IW # | Avg. NO on Screening |
|-----|-----------|---------|-----------|-----------------------------------|----------------|-----------------|-----------------|------------------------|--------------------|--------------|----------------------|
| 1 | 52410702 | 57 | 1.1 | 55.3 | 160.9 | 0.0 | 0.0 | 0.0 | 1.1 | 66 | 10.0 |
| 2 | 52410704 | 49 | 1.1 | 67.8 | 168.0 | 0.0 | 0.0 | 0.1 | 1.0 | 54 | 8.8 |
| 3 | 52410710 | 13 | 1.3 | 51.9 | 173.8 | 0.0 | 0.0 | 0.0 | 1.0 | 17 | 9.5 |
| 4 | 52410717 | 127 | 1.2 | 57.8 | 171.5 | 0.0 | 0.0 | 0.1 | 1.1 | 164 | 10.1 |
| 5 | 52410718 | 141 | 1.2 | 76.1 | 194.1 | 0.0 | 0.0 | 0.0 | 1.1 | 171 | 9.2 |
| 6 | 52410719 | 20 | 1.2 | 69.1 | 167.0 | 0.0 | 0.0 | 0.1 | 1.0 | 23 | 10.1 |
| 7 | 52410723 | 34 | 1.2 | 64.0 | 163.5 | 0.0 | 0.0 | 0.1 | 1.0 | 40 | 9.5 |
| 8 | 52410726 | 153 | 1.2 | 56.2 | 162.7 | 0.0 | 0.0 | 0.4 | 1.1 | 182 | 10.4 |
| 9 | 52410727 | 159 | 2.0 | 41.1 | 123.3 | 0.0 | 0.0 | 2.5 | 1.0 | 388 | 9.5 |
| 10 | 52410728 | 178 | 1.1 | 65.5 | 152.7 | 0.0 | 0.0 | 0.0 | 1.0 | 208 | 8.3 |
| 11 | 52410729 | 165 | 1.4 | 55.7 | 185.7 | 0.0 | 0.0 | 0.0 | 1.0 | 231 | 8.9 |
| 12 | 52410730 | 149 | 1.5 | 50.5 | 177.2 | 1.0 | 1.0 | 0.2 | 1.0 | 215 | 8.8 |
| 13 | 52410731 | 78 | 1.2 | 63.0 | 171.1 | 0.0 | 0.0 | 0.1 | 1.0 | 95 | 8.4 |
| 14 | 52410734 | 172 | 1.5 | 63.5 | 195.3 | 0.0 | 0.0 | 3.1 | 1.0 | 354 | 6.2 |
| 15 | 52410735 | 70 | 1.1 | 63.1 | 157.2 | 0.0 | 0.0 | 0.0 | 1.2 | 77 | 8.7 |
| 16 | 52410736 | 151 | 1.3 | 73.2 | 180.1 | 1.0 | 1.0 | 0.0 | 1.1 | 200 | 7.9 |
| 17 | 52410737 | 168 | 1.2 | 65.7 | 153.8 | 0.0 | 0.0 | 0.1 | 1.1 | 210 | 7.5 |
| 18 | 52410740 | 147 | 1.7 | 45.8 | 153.8 | 1.9 | 1.9 | 0.2 | 1.1 | 238 | 8.1 |
| 19 | 52410746 | 28 | 1.3 | 71.1 | 187.8 | 0.0 | 0.0 | 0.0 | 1.1 | 39 | 9.7 |
| 20 | 52410747 | 179 | 1.3 | 74.2 | 231.2 | 0.0 | 0.0 | 0.5 | 1.1 | 231 | 6.9 |
| 21 | 52410748 | 169 | 1.2 | 60.4 | 166.5 | 0.0 | 0.0 | 0.2 | 1.0 | 217 | 10.4 |
| 22 | 52410750 | 139 | 1.6 | 64.1 | 181.5 | 0.0 | 0.0 | 3.2 | 1.0 | 288 | 5.7 |
| **Grand Total** | | **2546** | **1.4** | **60.0** | **171.6** | **1.75** | **1.75** | **0.91** | **1.05** | **3708.0** | **8.37** |

# Nepal: Biomarker Collection

- Project also included collection of saliva samples from all respondents, for genotyping for mental disorders

- As part of sample management system, a web-based logging portal was used by data collection team and genetic lab in Kathmandu to track the current status of each saliva sample from field collection → shipment to genotyping institute in Boston, U.S.

# "WebLog"

# Kingdom of Saudi Arabia

- Again, survey of mental health prevalence

- Concerns about sensitivity and falsification

- Scripted interview verification (phone or face-to-face)

- Used SurveyTrak and questionnaire data to generate calculated "indicators" that assisted managers to target potential problems; **data driven assessment**

# QC Indicators by Type & Source of Error

| Source of Error | Single Occurrence Indicator | Cumulated Indicator |
|---|---|---|
| **Measurement** | • Any pause >= 10 minutes<br>• Any question read < 1 second<br>• An interview length < 30 minutes<br>• # of completed interviews >= 3 on the same day<br>• Failed verification | • Rate of verifications with discrepancy<br>• Rate of short path interviews<br>• Rate of no mental health disorders<br>• Short average interview length<br>• Rate of switching from ACASI to CAPI |
| **Coverage** | • Short travel time between two interviews on the same day<br>• Three interviews with a household member deleted from the roster<br>• Failed verification | • Rate of cases that are unable to verify<br>• (H) Rate of household with no eligible female/male |
| **Nonresponse** | | • Rate of saliva refusal<br>• Lowest average contact attempts per completed interview<br>• Low response rate |

20

# QC Summary Report

| Iwer ID | Count of Completed Screeners | Count of Completed Main Interviews | Not read Question | Failed Verification | Short IW | Long pause | Deleted HH member | Number of completed IWs/day | Short Time between IWs | Sum of Flagged Indicators |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 11 | 6 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 2 |
| 2 | 6 | 2 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 3 |
| 3 | 4 | 13 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 3 |
| 4 | 31 | 15 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 4 |
| 5 | 29 | 69 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 3 |
| 6 | 22 | 29 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 5 |
| 7 | 20 | 26 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 3 |
| 8 | 38 | 38 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 3 |
| 9 | 17 | 16 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 4 |
| 10 | 31 | 27 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 3 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| Etc.. | | | | | | | | | | |

# Example: Quick Read Drill Down

# Example: Endorsing Stem Q's

# Qatar

- Identification of **25** QC indicators from sample management system and questionnaire

- Development of QC indicator processor that applies indicator flags

- Charts and tables to use as visualization

There are 4 areas of QC Indicators - total of 25 Indicators

1.  Overall Visit
    - 1.1. Low Average Number of Visits
    - 1.2. Low Response Rate
    - 1.3. Too many completed interviews per day
    - 1.4. Short duration between completed interviews (in minutes)
2.  Respondent Listing, Eligibility, Selection, and Participation
    - 2.1. Low percentage of household with at least one eligible fe/male member
    - 2.2. Lowe average number of eligible fe/male household members listed
3.  Main Interview
    - 3.1. Short question field time - under 1 second
    - 3.2. Short interview length
    - 3.3. Short average interview length
    - 3.4. High Number of Negative Stem
    - 3.5. Low prevalence rate
    - 3.6. Long pause
4.  Study Components
    - 4.1. Low saliva given rate
    - 4.2. Low ACASI usage rate
    - 4.3. (VER) Low rate of household w phone number
    - 4.4. (FTF VER) High rate of verification w/unable-to-verify status (unable: cannot locate the household)
    - 4.5. High rate of verification w/inconsistent status (inconsistent: discrepant GPS distance)
    - 4.6. (FTF VER) High rate of verification w/inconsistent status (inconsistent: wrong household being interviewed)
    - 4.7. (FTF VER:) Frequent violation of fielding practice - first visit
    - 4.8. (FTF VER) Frequent violation of fielding practice - second visit
    - 4.9. (TEL VER: Demographic ) Failed
    - 4.10. (TEL VER: Mental Health) Failed
    - 4.11. (TEL VER: Demographic ) High rate of verification w/inconsistent outcome
    - 4.12. (TEL VER: Mental Health) High rate of verification w/inconsistent outcome
    - 4.13. (TEL VER) High rate of verification w/ unable-to-verify status (unable: cannot call/not home)

25

| Release | 3 | .T | | LEVEL 1 INDICATORS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Grand Total ** | 371 | 111 | 11 | 5 | 1 | 0 | 11 | 1 | 4 | 8 |
| Row Labels | D_Main_1001_1005_ADT | Sum_QC_Indicators | QC_Qtime_Main | QC_Short_Main_Sh | QC_Short_Main_LT | QC_Short_Main_LO | QC_Long_Paus | QC_Delete_HH | QC_Iw_Day | QC_Short_Tim_Bw_Iw |
| Mohammad, AbduAllah | 41 | 15 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 |
| Naser, AbdAlRahman | 40 | 14 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| AlTanani, AbdAlMajeed | 49 | 13 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 |
| Abdulqadir, Ameera | 37 | 11 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| Mohammad, Arwa | 30 | 8 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 |
| Omran, Sulaiman | 35 | 8 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| Yousuf, Amal | 36 | 8 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| AlTumani, Muna | 40 | 8 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| Fatah, Muna | 2 | 8 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| Salman, Hakima | | | | | | | | | | 1 |
| Mohammad, Alia | | | | | | | | | | 1 |
| Albarak, Ahmad | | | | | | | | | | 0 |
| Omran, Ahmad | | | | | | | | | | 0 |
| Abdelkader, Sabah | | | | | | | | | | 0 |
| Al-Mahzamah, Mohamad | | | | | | | | | | 0 |
| Eldali, Elaaf | | | | | | | | | | 0 |

## QC Indicator Summary

**Select Iwer (Device ID):**

| DeviceID | # CaseID | Total flags | Shortiw_Flag | Longiw_Flag | LongPause_Flag | QTimeShort_Flag | Prevalence_Flag | Endorse_Flag |
|---|---|---|---|---|---|---|---|---|
| SESRILP136 | 1 | 8 | | | 0 | 7 | 0 | 1 |
| SESRILP147 | 1 | 6 | | 1 | 0 | 3 | 1 | 1 |
| SESRILP169 | 2 | 6 | | 1 | 1 | 3 | 0 | 1 |
| SESRILP123 | 1 | 5 | 1 | | 0 | 3 | 0 | 1 |
| SESRILP133 | 1 | 5 | | | 0 | 3 | 1 | 1 |
| SESRILP130 | 1 | 5 | | 1 | 0 | 3 | 0 | 1 |
| SESRILP167 | 1 | 5 | | 1 | 0 | 3 | 0 | 1 |
| SESRILP171 | 2 | 5 | | 1 | 0 | 2 | 1 | 1 |
| SESRILP103 | 1 | 5 | 1 | | 0 | 2 | 1 | 1 |
| SESRILP112 | 1 | 4 | | | 0 | 3 | | 1 |
| SESRILP122 | 1 | 4 | | 1 | 0 | 2 | | 1 |
| SESRILP132 | 1 | 4 | | 1 | 0 | 2 | | 1 |
| SESRILP139 | 1 | 4 | | 1 | 1 | 1 | | 1 |
| SESRILP150 | 1 | 4 | | 1 | 0 | 2 | | 1 |
| SESRILP166 | 1 | 4 | | | 0 | 2 | 1 | 1 |
| SESRILP178 | 2 | 4 | | | 0 | 2 | 1 | 1 |
| SESRILP181 | 1 | 4 | | 1 | 1 | 1 | | 1 |
| SESRILP164 | 2 | 4 | | | 0 | 2 | 1 | 1 |
| SESRILP187 | 1 | 4 | | | 0 | 2 | 1 | 1 |
| SESRILP109 | 1 | 3 | | | 0 | 2 | 0 | 1 |
| SESRILP117 | 2 | 3 | 1 | | 0 | 1 | | 1 |
| SESRILP124 | 1 | 3 | | 1 | 0 | 1 | | 1 |
| SESRILP129 | 1 | 3 | | 1 | 0 | 1 | | 1 |
| SESRILP141 | 1 | 3 | | 1 | 0 | 1 | | 1 |
| SESRILP142 | 1 | 3 | | 1 | 0 | 1 | | 1 |
| SESRILP177 | 1 | 3 | | | 0 | 2 | 0 | 1 |
| SESRILP179 | 1 | 3 | | | 0 | 2 | | 1 |
| SESRILP180 | 1 | 3 | | | 0 | 2 | | 1 |
| SESRILP183 | 1 | 3 | | 1 | 0 | 1 | | 1 |
| Total | 95 | 234 | 8 | 19 | 3 | 115 | 8 | 81 |

**Survey Time(mins)**

| 33.73 | 0.04 | 1,429.77 |
|---|---|---|
| Average | Min | Max |

**Survey Questions Answered**

| 197.01 | 0 | 619 |
|---|---|---|
| Average | Min | Max |

**Interview Status**

2 10.37%
0
14.81%
4 74.81%

**Respondent Gender**

Male
66.74%
53.26%
Female

**Completed Interviews Per Day**

Jul 07, 12AM — Jul 07, 6AM — Jul 07, 12PM — Jul 07, 6PM — Jul 08, 12AM

# India

- National Data Innovation Centre (Gates Foundation)

- Using sample management and questionnaire paradata and data for all of the above, **plus** Computer Assisted Recorded Interview (CARI)

# CARI

- The CAPI system is used to trigger recording of a small number of questionnaire items

- Native laptop microphone is used, and CAPI captures interviewing screens

- Digital "movie" is encrypted and transmitted with survey data, and files can be re-played by supervisors in centralized office

- Interviewers given immediate feedback on adherence to standardized interviewing and project-specific protocols

# Conclusions

- Importance of project design and optimized use of technical systems
- Importance of interviewer training (certification) and on-going quality control
- Both sample management and questionnaire data are critical to detailed monitoring of production and quality
- Real-time monitoring of production allows you to identify problems and implement responsive and adaptive survey design changes
- Real-time monitoring of quality allows you to identify problems with specific interviewers and/or with the questionnaire or protocol
- Basic tools can be implemented with almost any technical system; more advanced tools are increasingly available with the recognition of the importance of collecting paradata and of making information available in an accessible way to local production teams

# References

- Mneimneh, Z., L. Lyberg, S. Sharma, M. Vyas, D. BalSathe, F. Malter, Y. Altwaijri; "Case studies on monitoring interviewer behavior in international and multinational surveys", in *Advances in Comparative Survey Methods: Multinational, Multiregional, and Multicultural Contexts (3MC);* Wiley (2019)

- Groves, R.M. "Research on survey data quality"; *Public Opinion Quarterly* 51: S156-S172 (1987)

- L. Lyberg and D. M. Stukel, "Quality assurance and quality control in cross-national comparative studies," in *Survey Methods in Multicultural, Multinational, and Multiregional Contexts*, Wiley (2010).