Computer-Assisted Personal Interviewing

TALIP KILIC
Living Standards Measurement Study
Poverty & Inequality Group
Development Research Group

ZURAB SAJAIA
Computational Tools Group
Development Research Group

Development Impact Evaluation

Field Coordinator Training
Washington, DC
April 22-25, 2013
Computer-Assisted Personal Interviewing (CAPI)

- **PAPI**: Paper-based personal interviewing, coupled with computer-assisted field-based data entry (CAFE) pioneered by the LSMS

- **CAPI**: Integration of interviewing & data entry through the use of a handheld device, preloaded with an electronic questionnaire

- Household (HH) surveys implemented on CAPI platform since the late’ 80’s in high- & middle-income countries, inc. the Netherlands, the UK, the US, Norway & Turkey

- Increasing number of applications in low-income setting in recent years
  - Mobile Phones, PDAs vs. **Netbooks & Tablet PCs**

- Need to balance urgency & quality, to emphasize value of experimentation & research
LSMS Experience

- LSMS operations marked by a gradual transition to CAPI
  - 2003 – CAPI survey experiment (Albania, CSProX, 200 HHs)
  - 2007 – CAPI survey (Ngara District, Tanzania, CWEST, ~500 HHs)
  - 2010 – KHDS (Kagera District, Tanzania, CWEST)
  - 2009-2014 – Uganda National Panel Survey
    - Supported by LSMS-ISA, implemented by UBoS
    - Partial transition to CAPI in 2010/11 (CWEST); CAPI transition completed in 2011/12 (CWEST & Surveybe); Next round 2013/14
  - 2011-2014 – Ethiopia Rural Socioeconomic Survey
    - Supported by LSMS-ISA, implemented by CSA
    - Partial transition to CAPI in 2011/12 for the Agriculture Questionnaire, in a subset of EAs (Surveybe); Next round in 2013/14
Hardware Options

General Features:

- 7-10” stylus-/touch-friendly screens
  - Rapid navigation across questionnaire
  - Several questions displayed at one time

- Camera, microphone, virtual keyboard & hand-writing recognition software

- 5-7 hours of (initial) battery life
  - Extended battery pack, external battery pack & daily charge of batteries recommended
  - Generators in low-electrification settings

- Multiple ports: Internet dongles, GPS units, external keyboards

Samsung Q1b Ultra (KHDS; UNPS) $650-700

Asus Eee PC T101MT (UNPS; ERSS) $450-500

Google Nexus 10 Tablet $400-600
Software Options

- Traditional DE software designed for transfer from paper questionnaire to computer
  - CAPI benefits better realized working with software packages designed for interactive interviewing

- CAPI software packages make up a small market, with varying degrees of cost effectiveness & strengths
  - Key players: *Blaise, CASES, CSProX, MMIC & Surveybe*


- CAPI Platform development underway in DECCT, in collaboration with the LSMS
Why CAPI?

- Expected gains in timeliness of data availability
  - Data entry, checking & exportation in one application
- Enhanced tools for in-field/remote management of teams under one roof
  - Headquarters & Team Leaders: Assigning work, tracking progress, immediate & comprehensive feedback
- Expected gains in scope & quality of data
  - Real time survey experiments/questionnaire randomization
  - GPS coordinates, time stamps, audio & video recording
  - Accommodation of non-linear/integrated questionnaires
  - Automated routing reduces the incidence of missing data
  - Data checking, reporting & revision facilities during the interview
  - Improvements in quantification of non-standard units
  - Instructions on questions, note taking facilities
Non-linear Navigation

We would like to make a complete list of Household members in the last 12 months including guests who slept here last night and those that left permanently.

<table>
<thead>
<tr>
<th>Wave1</th>
<th>Name</th>
<th>Gender</th>
<th>Residential Status</th>
<th>Age</th>
<th>Moved</th>
<th>Died</th>
<th>Demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mugalu Mohammed</td>
<td>Male [1]</td>
<td>Usual member present [1]</td>
<td>3</td>
<td>☐</td>
<td>☐</td>
<td></td>
</tr>
</tbody>
</table>

Tracking information
Automated Routing

T3ba Education Details

Education 1 | Education 2 | Education 3

Household ID: 1021000402

1. IS Gaddafi ANSWERING FOR HIMSELF OR HERSELF? Yes [1]

2. WHAT IS THE NAME OF THE PERSON RESPONDING FOR Gaddafi?

3. Can Gaddafi read and write with understanding in any language? Unable to read and write [1]


5. Why has Gaddafi not attended school? ,

6. What was the highest grade/class that Gaddafi completed?

7. What was the main reason that Gaddafi left school?

8. What grade/class was Gaddafi attending in the last complete
Consistency Checks

T3ca Health Details

Household ID 1021000402

Check

1. IS Mugalu ANSWERING FOR HIMSELF OR HERSELF? No [2]

2. WHAT IS THE NAME OF THE PERSON RESPONDING FOR Mugalu? Kimbowa Martin

3. During the past 30 days, did Mugalu suffer from any illness or injury? Yes [1]

4. For how many days did Mugalu suffer due to illness or injury during the past 30 days? 4

5. For how many days did Mugalu have to stop doing Mugalu’s usual activities due to illness or injury during the past 30 days? 12

6. Can you describe the symptoms that Mugalu primarily suffered due to Diarrhoea (acute) [1]

7. Can you describe the symptoms that Mugalu primarily suffered due to

8. Was anyone consulted (e.g. a doctor, nurse, pharmacist or traditional healer) for the major... Yes [1]
## Consistency Checks (Cont’d)

<table>
<thead>
<tr>
<th>Wave 1</th>
<th>Name</th>
<th>Age</th>
<th>Health Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mugalu Mohammed</td>
<td>3</td>
<td>![a. Health]</td>
</tr>
<tr>
<td>2</td>
<td>Gadaffi Ismail</td>
<td>25</td>
<td>![a. Health]</td>
</tr>
<tr>
<td>3</td>
<td>Kimbowa Martin</td>
<td>19</td>
<td>![a. Health]</td>
</tr>
</tbody>
</table>

**Household ID:** 1021000402

[Check]
Consistency Checks (Cont’d)

Household number: 1021000402

General comments

Question specific comments

Complete validation check

Update Visit Result ➤ V
Use of Media for Better Quantification

<table>
<thead>
<tr>
<th>Food consumption item</th>
<th>105 Sweet potatoes (fresh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household ID</td>
<td>1021000402</td>
</tr>
<tr>
<td>Over the last 7 days</td>
<td></td>
</tr>
</tbody>
</table>

1. How many days in the past 7 days was Sweet potatoes (fresh) consumed? [ ]

2. Unit of quantity [ ]

3. Amount of Sweet potatoes (fresh) consumed in the HOUSEHOLD over the last 7 days which came from PURCHASE... [ ]

4. Amount of Sweet potatoes (fresh) consumed AWAY FROM HOME over the last 7 days which came from PURCHASES... [ ]

5. Amount of Sweet potatoes (fresh) consumed over the last 7 days which came from HOME PRODUCTION? [ ]

6. Amount of Sweet potatoes (fresh) consumed over the last 7 days which came from GIFTS/IN-KIND sources? [ ]

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Use of Media for Better Quantification

Nomia Tin (500g)
Managing Expectations

- Data quality control principles in CAPI set-up no different than surveys based on PAPI with CAFE
- CAPI tools useful as much as enumerators & field supervisors take advantage of available facilities & act on inconsistencies
- Relative impact of CAPI on data quality: Open question, more experimentation & research necessary
  - Limited evidence on improved data quality with respect to a well-supervised survey based on PAPI with CAFE
Cost Implications

- CAPI generates (minimal) savings in printing costs & data entry
  - Savings increase with the complexity & frequency of survey

- Significant up-front costs in programming & hardware procurement
  - More cost-effective if machines are used in other survey operations
  - Need to budget for the possibility of hardware failure

- Increased need for more pilots ahead of the field work

- Transition into CAPI also driven by field work structure
  - Size of the enumerator corps may be prohibitively large

- Gradual transition to CAPI as part of the LSMS operations primarily underlined by demand for increased data quality & availability
Uganda National Panel Survey (UNPS) - CAPI Experience

- Teams quick to adapt, instrumental in training & knowledge sharing
- Required change in institutional thinking on HH survey: Greater up-front work (& costs) with respect to PAPI with CAFE
  - Prep of Wave I (PAPI) data uploaded onto Wave II (CAPI) application
  - Intra/inter-module consistency checks, in addition to range & default checks for missing values
  - Programming of rules on generation of household & individual identifiers for new additions to the sample
  - Training of UBoS Headquarters staff on case management suite
UNPS CAPI Experience (Cont’d)

• In-country procurement problems
  – Lags assoc. with operating within Government systems/unreliable suppliers
  – US procurement by the LSMS-ISA project: Not straightforward either
• Anti-virus software critical to maintaining the hardware integrity
• Application glitches even after piloting three times: Need for more intensive testing in comparison to PAPI with CAFE
• CAPI application platform based on multiple software packages: CWEST & CSPro (in 2010/11); CWEST & Surveybe (in 2011/12)
  • Dependence on the CWEST application developer for adjustments
  • Continued reliance on multiple software packages necessitated by lack of case management features on Surveybe
• Timely communication of bugs that might compromise the integrity of incoming data critical: No paper questionnaires to re-enter
UNPS CAPI Experience (Cont’d)

• Continuing improvements to the CAPI application on a rolling basis throughout the field work
• Even with internet dongles, slow internet speeds & lack of service in certain areas
  – Affects timely headquarters review of data sent from the field
  – Receipt of application updates by the survey teams not always timely
• Regular backup of interview files in the field & at the HQ crucial
• Still need a paper questionnaire for dissemination purposes: CAPI application dictionary is not more than a linear questionnaire report
Initially twofold objective:

- Inform internal decision making on the choice of surveys for upcoming surveys planned under the LSMS-ISA project, in Uganda, Ethiopia, and Nigeria
- Fill the gap in public knowledge on the relative performance of available software packages for the development of CAPI applications for multi-topic household surveys

- Peer-reviewed report, managed by the LSMS team, compiled by the IRIS Center at the University of Maryland, reviewed by software developers prior to release

Software packages screened as suitable for the development of CAPI applications for multi-topic household surveys & evaluated by the report include:

<table>
<thead>
<tr>
<th>Software</th>
<th>Developer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blaise</td>
<td>Westat &amp; Statistics Netherlands</td>
</tr>
<tr>
<td>CASES</td>
<td>Computer-Assisted Survey Methods Program at the UC Berkeley</td>
</tr>
<tr>
<td>CSProX</td>
<td>Serpro, S. A.</td>
</tr>
<tr>
<td>Entyware</td>
<td>Techneos</td>
</tr>
<tr>
<td>MMIC</td>
<td>RAND Labor and Population</td>
</tr>
<tr>
<td>Open Data Kit</td>
<td>Dept. of Computer Science &amp; Engineering at the Univ. of Washington</td>
</tr>
<tr>
<td>Pendragon Forms</td>
<td>Pendragon Software Corporation</td>
</tr>
<tr>
<td>Surveybe</td>
<td>Economic Development Initiatives</td>
</tr>
</tbody>
</table>
Comparative Assessment… (Cont’d)

- Structure of the report
  - Brief overview of each software package
  - Comparative assessment of each software package in 12 areas:

<table>
<thead>
<tr>
<th>Evaluation Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programming</td>
</tr>
<tr>
<td>Questionnaire Development</td>
</tr>
<tr>
<td>Questionnaire Implementation</td>
</tr>
<tr>
<td>Interface for Field Users</td>
</tr>
<tr>
<td>Questionnaire Navigation</td>
</tr>
<tr>
<td>Case Management</td>
</tr>
</tbody>
</table>

- Detailed evaluation of each software package, accompanied with full functionality check lists
Comparative Assessment… (Cont’d)

- No software package is an unequivocal frontrunner in all evaluation areas
- Missing across the landscape: A menu-driven development environment for novice users & a command line for more experienced users
- Positive relationship between quality/scope of documentation & proprietary nature of the software (MMIC, ODK vs. Blaise)
- Top contenders:
  - **Surveybe**: Ease to use (menu-based development environment) but lacks case management suite & only allowed for sequential workflow for questionnaire development at the time (under development)
  - **MMIC & Blaise**: Powerful & expansive in scope but steep learning curve (command line driven development environment) & high need for technical assistance
Sustainability of adoption relies on availability of a user-friendly, yet highly customizable, public solution

LSMS & Development Economics Computational Tools (DECCT) Unit supporting the development of a free CAPI software platform

Initially informed by LSMS field experience & comparative CAPI software assessment, recent pilot activities as well

Core interface components: **Designer** (for Survey Designers), **Survey Manager** (for Survey Management at the Headquarters & Team Leaders in the Field) & **Client** (for Interviewers)