Evaluating the Accuracy of Data Collection on Mobile Phones: A Study of Forms, SMS, and Voice

Somani Patnaik¹, Emma Brunskill¹, William Thies²

¹ Massachusetts Institute of Technology
² Microsoft Research India

ICTD 2009
Mobile Data Collection is in Style

• Especially in the developing world
  – Mobile banking
  – Microfinance
  – Healthcare
  – Environmental monitoring

• Benefits:
  – Faster
  – Cheaper
  – More accurate

Russia to introduce mobile phone voting in 2011

At the 2011 parliamentary elections, Russians will be able to cast their votes via their mobile phones, the Central Elections Commission has said.

No prior study of entry accuracy (on low-cost phones in developing world)
Data Collection on Mobile Phones

OpenROSA
FrontlineSMS Forms [Banks]
Nokia Data Gathering [Nokia]
RapidSMS [UNICEF]
MobileResearcher [Populi.net]
Cell-Life in South Africa [Fynn]
Jiva TeleDoc in India [UN Publications]
Pesinet in Mali [Balancing Act News]
Malaria monitoring in Kenya [Nokia News]
Voxiva Cell-PREVEN in Peru [Curioso et. al]
Data Collection on Mobile Phones

OpenROSA
FrontlineSMS Forms [Banks]
Nokia Data Gathering [Nokia]
RapidSMS [UNICEF]
MobileResearcher [Populi.net]
Cell-Life in South Africa [Fynn]
Jiva TeleDoc in India [UN Publications]
Pesinet in Mali [Balancing Act News]
Malaria monitoring in Kenya [Nokia News]
Voxiva Cell-PREVEN in Peru [Curioso et. al]

Data Collection on PDAs

SATELLIFE
EpiHandy
EpiSurveyor [Datadyne]
Infant health in Tanzania [Shrima et al.]
e-IMCI in Tanzania [DeRenzi et al.]
Respiratory health in Kenya [Diero et al.]
Tobacco survey in India [Gupta]
Ca:sh in India [Anantramanan et al.]
Malaria monitoring in Gambia [Forster et al.]
Clinical study in Gabon [Missinou et al.]
Tuberculosis records in Peru [Blaya et al.]
Sexual surveys in Peru [Bernabe-Ortiz et al.]
### Data Collection on Mobile Phones

- OpenROSA
- FrontlineSMS Forms [Banks]
- Nokia Data Gathering [Nokia]
- RapidSMS [UNICEF]
- MobileResearcher [Populi.net]
- Cell-Life in South Africa [Fynn]
- Jiva TeleDoc in India [UN Publications]
- Pesinet in Mali [Balancing Act News]
- Malaria monitoring in Kenya [Nokia News]
- Voxiva Cell-PREVEN in Peru [Curioso et al.]

### Data Collection on PDAs

- SATELLIFE
- EpiHandy
- EpiSurveyor [Datadyne]
- Infant health in Tanzania [Shrima et al.]
- e-IMCI in Tanzania [DeRenzi et al.]
- Respiratory health in Kenya [Diero et al.]
- Tobacco survey in India [Gupta]
- Ca:sh in India [Anantramanan et al.]

### Published Error Rates

- Malaria monitoring in Gambia [Forster et al.]
- Clinical study in Gabon [Missinou et al.]
- Tuberculosis records in Peru [Blaya et al.]
- Sexual surveys in Peru [Bernabe-Ortiz et al.]
<table>
<thead>
<tr>
<th>Data Collection on Mobile Phones</th>
<th>Data Collection on PDAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>OpenROSA</td>
<td>SATELLIFE</td>
</tr>
<tr>
<td>FrontlineSMS Forms [Banks]</td>
<td>EpiHandy</td>
</tr>
<tr>
<td>Nokia Data Gathering [Nokia]</td>
<td>EpiSurveyor [Datadyne]</td>
</tr>
<tr>
<td>RapidSMS [UNICEF]</td>
<td>Infant health in Tanzania [Shrima et al.]</td>
</tr>
<tr>
<td>MobileResearcher [Populi.net]</td>
<td>e-IMCI in Tanzania [DeRenzi et al.]</td>
</tr>
<tr>
<td>Jiva TeleDoc in India [UN Publications]</td>
<td>Tobacco survey in India [Gupta]</td>
</tr>
<tr>
<td>Pesinet in Mali [Balancing Act News]</td>
<td>Ca:sh in India [Anantramananan et al.]</td>
</tr>
<tr>
<td>Malaria monitoring in Kenya [Nokia News]</td>
<td></td>
</tr>
<tr>
<td>Voxiva Cell-PREVEN in Peru [Curioso et. al]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Published Error Rates</th>
<th>Published Error Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>None?</td>
<td>None?</td>
</tr>
<tr>
<td>CAM in India [Parikh et al.]</td>
<td>Malaria monitoring in Gambia [Forster et al.]</td>
</tr>
<tr>
<td>Clinical study in Gabon [Missinou et al.]</td>
<td>Clinical study in Gabon [Missinou et al.]</td>
</tr>
<tr>
<td>Tuberculosis records in Peru [Blaya et al.]</td>
<td>Tuberculosis records in Peru [Blaya et al.]</td>
</tr>
<tr>
<td>Sexual surveys in Peru [Bernabe-Ortiz et al.]</td>
<td>Sexual surveys in Peru [Bernabe-Ortiz et al.]</td>
</tr>
</tbody>
</table>
# Our Study

- Compared three interfaces for health data collection

<table>
<thead>
<tr>
<th>Electronic Forms</th>
<th>SMS</th>
<th>Live Operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 literate health workers &amp; hospital staff, Gujarat, India</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11. Patient’s Cough:</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Cough - Press 1</td>
</tr>
<tr>
<td>Rare Cough - Press 2</td>
</tr>
<tr>
<td>Mild Cough - Press 3</td>
</tr>
<tr>
<td>Heavy Cough - Press 4</td>
</tr>
<tr>
<td>Severe Cough - Press 5</td>
</tr>
<tr>
<td>(with blood)</td>
</tr>
</tbody>
</table>

— printed cue card—

| Error rate: | 4.2% | 4.5% | 0.45% |

Result caused partners to switch from forms to operator

- **Recommendations:**
  1. Caution needed in deploying critical apps w/ non-expert users
  2. A live operator can be accurate and cost-effective solution
Context: Rural Tuberculosis Treatment

• With local partners, working to improve tuberculosis treatment in rural Bihar, India

THE Prajnopaya Foundation 🌸 Innovators In Health

• Strategy: monitor patient symptoms remotely

Health worker uploads symptoms

Physician reviews, advises, schedules visits

• Data uploaded: 11 questions, every 2 weeks
  – Patient ID
  – Temperature
  – Weight
  – Cough (multiple choice)
  – Symptoms (yes / no)
Design Space:
Data Collection on Low-End Phones

Prompts

AUDIO

VISUAL

SMS
Electronic Forms
Voice-Activated Forms
Interactive Voice Response
Spoken Dialog
Live Operator

DATA ENTRY

less interactive
more interactive
less interactive
more interactive

TYPED
SPOKEN
Design Space: Data Collection on Low-End Phones

- **Prompts**
  - AUDIO
  - VISUAL

- **Data Entry**
  - TYPED
  - SPOKEN

- **Methods**
  - SMS
  - Electronic Forms
  - Live Operator

- **Interactivity**
  - less interactive
  - more interactive
1. SMS Interface

- **Pro:**
  + Potentially cheapest

- **Con:**
  - Easiest to fake visits
  - Least reliable

11. Enter the Patient’s Cough

<table>
<thead>
<tr>
<th>Condition</th>
<th>Press</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Cough</td>
<td>1</td>
</tr>
<tr>
<td>Rare Cough</td>
<td>2</td>
</tr>
<tr>
<td>Mild Cough</td>
<td>3</td>
</tr>
<tr>
<td>Heavy Cough</td>
<td>4</td>
</tr>
<tr>
<td>Severe Cough (with blood)</td>
<td>5</td>
</tr>
</tbody>
</table>

21. Check Yourself

Your finished message should be formatted similarly to the following:

10 372 62 68 4 1030007
2. Electronic Forms Interface

• **Pro:**
  + Arguably more user friendly than SMS

• **Con:**
  – Expensive handset
3. Live Operator Interface

- **Pro:**
  - Most flexible Q&A
  - No literacy required
  - Hard to fake visits

- **Con:**
  - Cost of operator
  - Potentially slower

---

Patient: "Are you having night sweats?"
Health Worker: "Is the patient having night sweats?"
Operator: "No, she isn’t.”

Patient: "No, I’m not.”
Study Participants

• 13 health workers and hospital staff (Gujarat, India)

<table>
<thead>
<tr>
<th></th>
<th>Age (Median)</th>
<th>Education</th>
<th>Cell Phone Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health workers (6)</td>
<td>23</td>
<td>10(^\text{th}) – 12(^\text{th})</td>
<td>Had used phone</td>
</tr>
<tr>
<td>Hospital staff (7)</td>
<td>30</td>
<td>12(^\text{th}) – D. Pharm.</td>
<td>Owned phone</td>
</tr>
</tbody>
</table>

• Within-subjects design

• Training standard: two error-free reports on each interface
  – Health workers: big groups, 6-8 hours
  – Hospital staff: small groups, 1-2 hours
Results

Append to current SMS:

11. Patient’s Cough:
   - No Cough - Press 1
   - Rare Cough - Press 2
   - Mild Cough - Press 3
   - Heavy Cough - Press 4
   - Severe Cough (with blood) - Press 5

--- printed cue card ---

<table>
<thead>
<tr>
<th>Electronic Forms</th>
<th>SMS</th>
<th>Live Operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error rate</td>
<td>4.2% (12/286)</td>
<td>4.5% (13/286)</td>
</tr>
</tbody>
</table>
## Results

<table>
<thead>
<tr>
<th></th>
<th>Health workers</th>
<th>Hospital staff</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Electronic Forms</strong></td>
<td>7.6%</td>
<td>1.3%</td>
</tr>
<tr>
<td><strong>SMS</strong></td>
<td>6.1%</td>
<td>3.2%</td>
</tr>
<tr>
<td><strong>Live Operator</strong></td>
<td>1.5%</td>
<td>0%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Error rate (errors / entries)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Electronic Forms</strong></td>
<td>4.2% (12/286)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SMS</strong></td>
<td>4.5% (13/286)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Live Operator</strong></td>
<td>0.45% (1/220)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sources of Error

- Multiple Choice (SMS)
- Multiple Choice (Forms)
- Numeric
Sources of Error

Usability Barriers
- small keys
- correcting mistakes
- decimal point

<table>
<thead>
<tr>
<th>Correct</th>
<th>Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>54</td>
<td>45</td>
</tr>
<tr>
<td>62</td>
<td>826</td>
</tr>
<tr>
<td>62</td>
<td>empty</td>
</tr>
<tr>
<td>68</td>
<td>67</td>
</tr>
<tr>
<td>68</td>
<td>93</td>
</tr>
<tr>
<td>69</td>
<td>59</td>
</tr>
<tr>
<td>98.5</td>
<td>98</td>
</tr>
<tr>
<td>98.7</td>
<td>98.687</td>
</tr>
<tr>
<td>100.2</td>
<td>100.0</td>
</tr>
<tr>
<td>100.3</td>
<td>103</td>
</tr>
<tr>
<td>“1003”</td>
<td>103</td>
</tr>
<tr>
<td>100.8</td>
<td>108</td>
</tr>
</tbody>
</table>
Sources of Error

Usability Barriers
- small keys
- correcting mistakes
- decimal point
- scrolling / selection

<table>
<thead>
<tr>
<th>Correct</th>
<th>Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
<td>None</td>
</tr>
<tr>
<td>Heavy</td>
<td>Mild</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
**Sources of Error**

**Usability Barriers**
- small keys
- correcting mistakes
- decimal point
- scrolling / selection
- SMS encoding

**Correct** | **Incorrect**
---|---
“1” (none) | “0” (disallowed)
“1” (none) | “0” (disallowed)
“1” (none) | “0” (disallowed)
“3” (mild) | “0” (disallowed)
“5” (severe) | *empty*
“6” (A. Khanna) | “5” (A. Kumar)
“7” (A. Kapoor) | “1” (A. Khan)
“6” | “2”
“0000007” | “000007”
Sources of Error

Usability Barriers
- small keys
- correcting mistakes
- decimal point
- scrolling / selection
- SMS encoding

Detectable Errors
- Multiple Choice (SMS)
- Numeric
- Multiple Choice (Forms)
## Cost Comparison

<table>
<thead>
<tr>
<th></th>
<th>SMS</th>
<th>Forms</th>
<th>Live Operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost per interview</td>
<td>$C_S$</td>
<td>$C_S$</td>
<td>$(C_V + C_O) \times T$</td>
</tr>
</tbody>
</table>

**Program variables**  
$T$ time spent per interview

**Cost variables**  
$C_S$ cost of an SMS  
$C_V$ cost of a voice minute  
$C_O$ cost of an operator minute
Cost Comparison

<table>
<thead>
<tr>
<th></th>
<th>SMS</th>
<th>Forms</th>
<th>Live Operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost per interview</td>
<td>$0.03</td>
<td>$0.03</td>
<td>$0.06 T</td>
</tr>
</tbody>
</table>

**Program variables**

- $T$ time spent per interview

**Cost variables in Bihar, India**

- $0.03$ cost of an SMS
- $0.02$ cost of a voice minute
- $0.04$ cost of an operator minute
# Cost Comparison

<table>
<thead>
<tr>
<th></th>
<th>SMS</th>
<th>Forms</th>
<th>Live Operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost per interview</td>
<td>$0.03</td>
<td>$0.03</td>
<td>$0.06 T</td>
</tr>
</tbody>
</table>

Break-even call: 30 seconds

**Program variables**
- $T$ time spent per interview

**Cost variables in Bihar, India**
- $0.03$ cost of an SMS
- $0.02$ cost of a voice minute
- $0.04$ cost of an operator minute
# Cost Comparison (TB Program)

<table>
<thead>
<tr>
<th></th>
<th>SMS</th>
<th>Forms</th>
<th>Live Operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost per interview</td>
<td>$0.03</td>
<td>$0.03</td>
<td>$0.15</td>
</tr>
<tr>
<td>Cost per phone</td>
<td>$25</td>
<td>$50</td>
<td>$25</td>
</tr>
<tr>
<td>Total cost</td>
<td>$29</td>
<td>$54</td>
<td>$43</td>
</tr>
</tbody>
</table>

**Program variables**
- **2.5 min** time spent per interview
- **120** number of interviews for duration of program

**Cost variables in Bihar, India**
- **$0.03** cost of an SMS
- **$0.02** cost of a voice minute
- **$0.04** cost of an operator minute

SMS < Live Operator < Forms
## Cost Comparison (Microfinance)

<table>
<thead>
<tr>
<th></th>
<th>SMS</th>
<th>Forms</th>
<th>Live Operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost per interview</td>
<td>$0.03</td>
<td>$0.03</td>
<td>$0.60</td>
</tr>
<tr>
<td>Cost per phone</td>
<td>$25</td>
<td>$50</td>
<td>$25</td>
</tr>
<tr>
<td>Total cost</td>
<td>$40</td>
<td>$65</td>
<td>$325</td>
</tr>
</tbody>
</table>

Microfinance: Operator is 5x more expensive than Forms

### Program variables
- **10 min**: time spent per interview
- **500**: number of interviews for duration of program

### Cost variables in Bihar, India
- **$0.03**: cost of an SMS
- **$0.02**: cost of a voice minute
- **$0.04**: cost of an operator minute
The Case for Live Operators

• Our proposition:
  Operators are under-utilized for mobile data collection

• Benefits:
  – Lowest error rate
  – Less education and training needed
  – Most flexible interface

• Challenges:
  – Servicing multiple callers
Related Work

• **Personal digital assistants (PDAs) for mobile health**
  – 8+ hours training, educated workers: 0.1% - 1.7% error rates [Forster et al., 1991] [Missinou et al., 2005] [Blaya & Fraser, 2006]
  – 2-3 minutes training, uneducated workers: 14% error rate [Bernabe-Ortiz et al., 2008]
  – In developed world: mixed results vs. paper forms [Lane et al., 2006]

• **Richer interfaces**
  – CAM: <1% error rates via camera phone [Parikh et al.]
  – Speech [Patel et al., 2009] [Sherwani et al. 2009] [Grover et al.] [ … ]
  – Interfaces for low-literate users [Medhi et al.]
Conclusions

• Accuracy of mobile data collection demands attention
  – We measured 5% error rates for those lacking experience

• There exist cases where a live operator makes sense
  – Error rates shrunk to 0.5%
  – Can be cost effective, esp. for short calls or infrequent visits

• Our study has limitations
  – Small sample size
  – Varied education, phone experience, training of participants

• Future work
  – Distinguish factors responsible for error rates
  – Compare to paper forms, Interactive Voice Response