Large-Scale Survey Interviewing Following

the 2008 WenChuan Earthquake

Zhang Huafeng and Jon Pedersen

International Blaise Users Conference (IBUC)
Baltimore, USA
Two household and community surveys
  - Post Wenchuan Earthquake Rapid Needs Assessment Survey
    - July 2008
    - Interviewed 3652 households
  - Restoration and reconstruction survey in the Wenchuan earthquake area
    - July 2009
    - Interviewed 4014 households

Both surveys were:
  - Large-scale computer-assisted personal interviewing (CAPI) surveys
  - To assist the Chinese government on the provincial and national level in early and mid-term recovery planning

First survey operated on a very strict timeline to meet the government’s schedule of developing General Plan
CAPI with Blaise

- **CAPI was the only choice due to the strict timeline for the first survey**
  - Only 10 days for field work, data work, and report writing
  - 80 Asus Eee-pc sub-notebook computers running Windows XP and using Blaise 4.8 as data entry tool
  - Saved time needed for data entry and allowed for continuous tabulation and quality control

- **Blaise was chosen because**
  - Blaise was reported by other researchers to be a useful and powerful environment for CAPI data entry application
  - Previous experiences with Blaise by the survey team
  - Standard blaise could be used with sub-notebook computer running Window XP
  - Detail data entry formatting is not needed
Surveys:

- Two stage cluster sample
- 144 clusters
- Stratified by type of area
- Reported by
  - Very seriously affected
  - Seriously affected
  - Camps
## Allocation of the sample

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Population size</th>
<th>Number of PSUs</th>
<th>Selected PSUs Reporting domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very seriously affected area</td>
<td>3,666,220</td>
<td>3,038</td>
<td>110 Very seriously affected</td>
</tr>
<tr>
<td>Seriously affected area</td>
<td>7,976,816</td>
<td>5,365</td>
<td>40 Seriously affected</td>
</tr>
<tr>
<td>Margin of seriously affected area</td>
<td>1,437,349</td>
<td>1660</td>
<td>10 Seriously affected</td>
</tr>
<tr>
<td>Deyang city camps</td>
<td>142,943</td>
<td>166</td>
<td>5 Camps</td>
</tr>
<tr>
<td>Mianyang camps</td>
<td>182,390</td>
<td>18</td>
<td>5 Camps</td>
</tr>
<tr>
<td>Guangyuan camps</td>
<td>42,244</td>
<td>19</td>
<td>5 Camps</td>
</tr>
<tr>
<td>Chengdu camps</td>
<td>103,905</td>
<td>15</td>
<td>5 Camps</td>
</tr>
</tbody>
</table>
• **Household**
  - Housing, infrastructure and amenities
  - Mortality in the earthquake and characteristics of the deceased
  - Basic demographic information including age, gender, ethnicity, marital status
  - Education, current enrollment and attendance
  - Health situation including injuries due to the earthquake
  - Work and relation to the labour market before and after the earthquake
  - Household economy, economic activities and agriculture
  - Economic support in the aftermath of the earthquake
  - Migration and plans to move
  - Social network

• **Randomly Selected Individual**
  - A 12 item psychological distress scale, CHQ-12 (Yang, Huang, & Wu, 2003)
  - Participation in rescue activities
  - Trust in persons and institutions
  - Attitudes to various forms of assistance to earthquake victims
  - Satisfaction with services/assistance
Advantages
Advantages compared to PDAs

• PDAs have successfully been employed in many surveys, but available software for PDAs are:
  – EITHER difficult to use (Handbase)
  – OR expensive (CSPRO-X)
  – OR too simple to handle a complex questionnaire (Pendragon Forms)

• Note-book running Windows-XP and Blaise are much better supported in the internet user community than PDAs
Dramatically reduced the time needed for the survey

• To reduce the survey time is critical in early stages of a humanitarian emergency
  – No time needed for paper transportation, data entry or cleaning
  – Reduce the interview time
  – Higher technical requirement of field operation
  – Simplify survey management (questionnaire management, transportation and field editing)

• Field work process for the first survey
  – 400-500 interviews daily, 5.2 interviews per interviewer per day

![Graph showing completed and sampled interviews over time]
Technical benefits and advantages

- Facilitate the interviewers to navigate the lengthy questionnaire

- Automatic tracking of the questions to be asked, and avoid skipping and missing errors

- Internal range and consistency checks, saved training, field editing and data cleaning time

- Improve random selection of respondents within the household
  - Question: Turned out to be a challenge for the programmer, due to the reevaluation of the calculated variables upon every change
Quality assurance during field work

• Possible to evaluate the performance and quality of the field work on the daily bases

• Possible to record the information about the interview process, such as interview time, recording sound and etc.
  – However, storing data entry time in variables in Blaise is subject to similar problems as the random number generation – namely reevaluation upon change

• Possible to incorporate the notes and interviewer manual into the blaise data entry program
Disadvantages
Hardware limitation and challenges in the field

• **Battery life**
  – Problem with the areas without electricity supply right after the earthquake
  – Possible to solve with the improvement of technology

• **Data loss due to the hardware failures or human errors**
  – Extra work with data management and equipment maintenance
  – Data automatic saved every minute with Blaise
  – Data backup during transfer
  – One technician available all times by phone
  – Immediate replacement of crashed computer
  – A few paper questionnaires carried by each interviewer

• **Screen readability**
  – Under bright sunlight
  – Small screen
    o Blaise is flexible in adjusting the font size
    o Large font is problematic for the question with many answers
Programming challenges

• Parallel development of the Blaise data entry program and the questionnaire

• Demand good organization and cooperation between the programmer and questionnaire designer

• Solve the bugs shown up during field work
  – Teams were sent to the same area and spent nights in the same place in the first two days, so that updating program was possible
  – The strict rules Blaise enforces for updating data entry application during the field work are quite beneficial here
• While double entry could minimize mis-keying errors in PAPI surveys, CAPI surveys are prone to introduce more undetectable mis-keying errors

• Panel surveys made it possible to detect the mis-keying errors
  – Interviewers committed more mis-keying errors when under pressure
    o 365 interviews per day (304 in the second stage) in the first survey, 1.4% mis-typed gender, 1.3% mis-typed residency status
    o 243 interviews per day in the second survey, 0.5% mis-typed gender, 0.4% mis-typed residency status

• Strategies for minimizing undetectable mis-keying errors
  – Assign values bit farther away from each other
  – Increase font size and add extra space between the answers
  – Avoid using auto-skip in Blaise
Mis-keying errors
Mis-keying of large numbers

- **Large value mis-keyings detected during training**
  - Large values such as income, fees and cost
  - Downward bias for large numbers and upward bias for small numbers
  - Large value mis-keys occur more frequently and affect data quality more seriously

- **Strategies for minimizing large value mis-keyings**
  - Warning box with values translated into Chinese characters

<table>
<thead>
<tr>
<th></th>
<th>Percentage of trainees who answered correctly</th>
<th>Correct answer (Yuan)</th>
<th>Mean of all the answers (Yuan)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family business income</td>
<td>62.3 %</td>
<td>120,000</td>
<td>86,589</td>
</tr>
<tr>
<td>Living assistance from government</td>
<td>81.8 %</td>
<td>2,700</td>
<td>8,400</td>
</tr>
<tr>
<td>Personal income in the past one year</td>
<td>79.5 %</td>
<td>14,400</td>
<td>20,209</td>
</tr>
<tr>
<td>Annual salary in the past one year</td>
<td>77.6 %</td>
<td>120,000</td>
<td>99,342</td>
</tr>
<tr>
<td>Current monthly salary</td>
<td>96.1 %</td>
<td>10,000</td>
<td>9,776</td>
</tr>
<tr>
<td>School fee in the past one year</td>
<td>61.0 %</td>
<td>5,000</td>
<td>5,319</td>
</tr>
</tbody>
</table>
Conclusion

- **Netbook computer served as an excellent platform for survey**
  - Lighter
  - Easier to use
  - More versatile than PDAs

- **Blaise proved to be an effective tool for developing data entry application rapidly for CAPI surveys in emergency settings**
  - Flexibility of Blaise programming language
  - Automatic handling of scree layout
  - Concurrent questionnaire development and programming possible
  - Type control on the data structure facilitate rapid development
• Thank you!

• zhu@fafo.no & jpe@fafo.no