Interviewer identity & learning effects: Sources of variation in reported social networks


Harvard T.H. Chan School of Public Health
University of the Witwatersrand

Health and Aging in Africa: Longitudinal Studies of INDEPTH Communities (HAALSI)

7 April 2016
Newport Beach, CA
Motivation

Our goal

1. To elicit measures of social connections
2. To conduct valid inference relating to social connections

(Some of) Our concerns

1. Variation in comprehension of questions by respondent
2. Unintentional under-reporting: recall bias
3. Intentional under-reporting: burden reduction; social desirability

But don’t forget...
Motivation

It takes two to interview
Motivation

Sources of interviewer influence

1. Variation in comprehension of questions
2. Unintentional variation in effort
   - Interviewer characteristics: age, gender, language
   - Dyad characteristics: differences in age, gender, language
   - Idiosyncratic variability
3. Intentional variation
   - Normative network size
   - Normative length of interview
   - Learning effects emerge over time

Marsden 2003; Van der Zouwen & Van Tilburg 2001
HAALSI study

Health and Aging in Africa: Longitudinal Studies of INDEPTCH communities

- Baseline for a longitudinal cohort to study aging & health
- In rural South Africa
HAALSI study site
HAALSI study site
HAALSI study

Health and Aging in Africa: Longitudinal Studies of INDEPTH communities

- Baseline for a longitudinal cohort to study aging & health
- In rural South Africa
- Resident adults aged 40 and above
- Random sample of 49% of all eligible individuals
  - 5059 valid responses (80.5% response rate)
- Comprehensive interviews on health and social wellbeing
  - Approximately 3 hours to complete
- Conducted between November 2014 and November 2015
Social network module

Single name generator

"Please tell me the names of 6 adults with whom you have been in communication either in person or by phone or by internet in the past 6 months, starting with the person who is most important to you for any reason"
Social network module

Multiple name interpreters

For each person, over the past 6 months, how often you:

- interacted in person, by phone, SMS, email or the internet
- received emotional, physical, informational, financial support
- physical fought, verbally argued or were criticized by

Additionally we asked about:

- Alter age, gender, relationship and place of residence
- Alter-alter tie strength
Research questions

1. Do we see variation according to interviewer?

2. Who are the alters we do not see?

3. Is variation by interviewer due to respondent characteristics?
## Results: interviewers

<table>
<thead>
<tr>
<th>Sex</th>
<th>Interviewers</th>
<th>%</th>
<th>Respondents</th>
<th>%</th>
<th>Mean Alters</th>
<th>95% CI</th>
<th>Chi sq.</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>7</td>
<td>24.1%</td>
<td>1519</td>
<td>30.0%</td>
<td>3.18</td>
<td>[3.09 - 3.27]</td>
<td></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Female</td>
<td>20</td>
<td>69.0%</td>
<td>3540</td>
<td>70.0%</td>
<td>3.03</td>
<td>[2.97 - 3.08]</td>
<td>17.85</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>Interviewers</th>
<th>%</th>
<th>Respondents</th>
<th>%</th>
<th>Mean Alters</th>
<th>95% CI</th>
<th>Chi sq.</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-29</td>
<td>17</td>
<td>58.6%</td>
<td>3672</td>
<td>72.6%</td>
<td>3.15</td>
<td>[3.10 - 3.21]</td>
<td></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>30-39</td>
<td>7</td>
<td>24.1%</td>
<td>1056</td>
<td>20.9%</td>
<td>3.15</td>
<td>[3.04 - 3.25]</td>
<td></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>40-49</td>
<td>3</td>
<td>10.3%</td>
<td>331</td>
<td>6.5%</td>
<td>1.95</td>
<td>[1.80 - 2.11]</td>
<td>159.55</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experience</th>
<th>Interviewers</th>
<th>%</th>
<th>Respondents</th>
<th>%</th>
<th>Mean Alters</th>
<th>95% CI</th>
<th>Chi sq.</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;200 interviews</td>
<td>13</td>
<td>44.8%</td>
<td>1284</td>
<td>25.4%</td>
<td>2.32</td>
<td>[2.22 - 2.41]</td>
<td></td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>≥ 200 interviews</td>
<td>14</td>
<td>48.3%</td>
<td>3775</td>
<td>74.6%</td>
<td>3.29</td>
<td>[3.23 - 3.34]</td>
<td>323.18</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Total: 27 | 5059
Results: survey timing

![Graph showing the mean number of contacts reported across different months, with the number of interviews indicated by different symbols.](image-url)
Results: interviewer

Guy Harling

Interviewer effects

Sunbelt XXXVI
Results: tie intensity

Guy Harling
Interviewer effects
Sunbelt XXXVI
Results: tie location

 Relatives

 Non-relatives
## Results

### Mixed effects Poisson Regression models

<table>
<thead>
<tr>
<th></th>
<th>Null</th>
<th>Months</th>
<th>Villages</th>
<th>Respondent</th>
<th>Interviewer</th>
<th>Dyad</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month of interview</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Respondent: Village of residence</td>
<td>0.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondent: Age decade (female)</td>
<td>0.03</td>
<td>0.02</td>
<td>0.35</td>
<td>0.02</td>
<td></td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Respondent: Age decade (male)</td>
<td>0.02</td>
<td>0.01</td>
<td>0.27</td>
<td>0.01</td>
<td></td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Respondent: Education</td>
<td>0.04</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td></td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Respondent: Country of origin</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondent: Marital status</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondent: Household size</td>
<td>0.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondent: Employment status</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondent: Household wealth</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interviewer: Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interviewer: Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dyad: Gender homophily</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.54</td>
</tr>
<tr>
<td>Dyad: Age difference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.81</td>
</tr>
<tr>
<td>AIC</td>
<td>18,346.5</td>
<td>17,805.4</td>
<td>17,830.3</td>
<td>17,191.8</td>
<td>17,195.2</td>
<td>17,200.2</td>
<td>17,192.1</td>
</tr>
<tr>
<td>Interviewer variance</td>
<td>0.12</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>[0.05 - 0.18]</td>
<td>[0.03 - 0.10]</td>
<td>[0.03 - 0.10]</td>
<td>[0.02 - 0.08]</td>
<td>[0.02 - 0.08]</td>
<td>[0.02 - 0.08]</td>
<td>[0.02 - 0.09]</td>
</tr>
</tbody>
</table>

Statistics are Wald tests across all categories of each independent variable.
Results: survey timing, adjusted
Key findings

1. Large changes in number of alters during a cross-sectional study with random interviewer assignment

2. Independent interviewer identity and time period effects

3. Not explainable by interviewer, respondent or dyad characteristics

4. Re-training may have led to a significant rise in alter numbers
What to do?

1. Less flexibility: fix number of alters required
2. More training: improve fieldworker comprehension
3. More monitoring: and feedback results to fieldworkers
Acknowledgements

- This work is supported by P01-AG-041710 from NIA (PI Lisa Berkman)

- This work is part of the HAALSI study: http://www.hsph.harvard.edu/population-development/research-focal-areas/major-projects/haalsi/

You can reach me at gharling@hsph.harvard.edu & @harlingg