The Total Survey Error Paradigm and Challenges to its Application in the Arab World

Kathrin Thomas\textsuperscript{1}  Michael Robbins\textsuperscript{1,2}

\textsuperscript{1} Princeton University
\textsuperscript{2} University of Michigan

5 June 2018
"Honest discussions about the extent, sources of, and solutions for quality issues [...] are needed. Regionally specific issues [...] raise concerns about social desirability and underscore the need for methodological research. Technological advances [...] offer possibilities for real-time monitoring and [...] research [...]. Yet apart from a handful of studies on interviewer effects, anchoring vignettes, and a few other topics, **almost no research systematically assesses the impact of the survey methods used on data quality in the Arab world.**"

(Benstead, 2018)
Is the Arab survey world just a few decades behind?

- Surveys and polling fairly new: beginning in 1980s, but real interest since the Arab uprising 2011 (Benstead, 2018)
- While core principles of Total Survey Error (TSE) are followed, holistic concept not (yet) recognised
- Survey practice in the Arab world may also make application of TSE more difficult
The TSE

(Groves & Lyberg, 2010)
Representation

- Potentially poor coverage due to past or ongoing security treats
- Absence of transparent sampling frames and/or outdated censuses
- Limited attention given to nonresponse and incomplete files
- Data fabrication, i.e, "Intentional departure from [...] sampling procedures by any member of the survey project" (Robbins, 2018)
Measurement

- Translation: Arabic ≠ Arabic ≠ Arabic
- Interviewer effects on response or item nonresponse (e.g., Benstead and Maluche, 2014; Benstead, 2014a, 2014b; Blaydes and Gillum, 2013; Koker, 2009)
- Third party presence (Benstead, 2014)
- Data fabrication, i.e., *intentionally* deviating from instructions; *intentionally* misrecording the answer given by a respondent; or *intentionally* creating data (Robbins, 2018)
The TSE...again

(Groves & Lyberg, 2010)
Data and Methods

- Arab Barometer III and IV
- Representation: Contact attempts, data fabrication
- Measurement: INR and response bias due to interviewers and third party presence
Interviewer Gender

Example Jordan

- Interviewer gender not included in the AB IV, but interviewer names available...in Arabic
- Three coders (2 Arabic speakers, 1 non-Arabic speaker) coded names into gender
  - $\kappa$ varied from 0.37 to 0.70
  - Non-Arabic speaker seemed to be least reliable; but also the Arabic speakers had discrepancies
  - Major challenges unisex names and context
Interviewer Gender

AB IV

- Interviewers in the Arab region predominantly female (AB IV: 72 percent)

Source: Arab Barometer Wave IV
Interviewer Religiosity

Approx. three quarters of the female interviewers wore a hijab during the survey interview (77.0 percent)

<table>
<thead>
<tr>
<th>Country</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algeria</td>
<td>N/A</td>
</tr>
<tr>
<td>Egypt</td>
<td>100</td>
</tr>
<tr>
<td>Jordan</td>
<td>96.3</td>
</tr>
<tr>
<td>Lebanon</td>
<td>28.5</td>
</tr>
<tr>
<td>Morocco</td>
<td>N/A</td>
</tr>
<tr>
<td>Palestine</td>
<td>94.2</td>
</tr>
<tr>
<td>Tunisia</td>
<td>29.1</td>
</tr>
</tbody>
</table>

**Table:** AB IV Proportion of Female Interviewers wearing a Hijab
Almost every other interview was attended by a third person (44.6 percent)
Little information about any of these aspects of the TSE

- Coverage problematic, esp. in less stable countries
  - e.g., field report for Iraq AB III indicates that only "more than 66 percent of the population of Iraq" are represented

- Sampling error problematic, esp. due to lack of transparent information

- No collection/no provision of UNR files

- Response rate largely AAPOR standard, but not for all countries
  - e.g., field report for AB IV Palestine suggest 91 percent response rate (response/response+refusal)
Example Tunisia

- Up to 8 contact attempts before UNR is logged
- On average, female interviewers need approximately one fewer contact attempt to interview compared to male interviewers
- T-test reveals statistically significant difference in contact attempts across interviewer gender ($t=2.7; p<0.01$)
Representation: Suspected Data Fabrication

Example 1

Source: Arab Barometer Wave III

Source: Arab Barometer Wave IV
Representation: Suspected Data Fabrication

Example 2

Source: Arab Barometer Wave III

Source: Arab Barometer Wave IV
### Table: Suspected Data Fabrication AB III and AB IV

<table>
<thead>
<tr>
<th></th>
<th>AB III</th>
<th>AB IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 percent match</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>95 percent match</td>
<td>161</td>
<td>0</td>
</tr>
<tr>
<td>90 percent match</td>
<td>336</td>
<td>4</td>
</tr>
<tr>
<td>85 percent match</td>
<td>565</td>
<td>18</td>
</tr>
<tr>
<td>Total no. of suspicious cases</td>
<td>1,066</td>
<td>22</td>
</tr>
<tr>
<td>Total no. of observations in survey</td>
<td>14,809</td>
<td>9,000</td>
</tr>
</tbody>
</table>
Representation: Implications

- Keep more detailed records of sampling procedures and frames
- Request UNR files by contract
- Request calculation of AAPOR RR by contract
- Development of a live data checking tool that flags up suspicious cases to prevent fabrication
Measurement: Interview Length

Example Tunisia

- Shortest interview just over 30 minutes; longest interview almost 2 hours
- Interviews conducted by female interviewers tend to be longer
  - T-test reveals statistically significant difference in total interview length across interviewer gender (t=2.7; p<0.01)
Overall, quite high INR (16.7 percent), substantive country variation.

Source: Arab Barometer Wave IV
Measurement: Item Nonresponse: Americans good people? (2)

- **Interviewer Gender**
  - Male interviewers appear to have higher INR ($\chi^2=22.7; p<0.01$) – all AB IV countries, but Algeria

- **Third Person Presence**
  - Third person presence appears to increase INR ($\chi^2=17.2; p<0.01$) – esp. Algeria, Morocco, and Tunisia
Measurement: Interviewer Gender & Women’s Rights (1)

- While female interviewers seem to evoke more agreement with the question whether women could become prime minister/president ($\chi^2=67.8$, $p<0.01$) – esp. Jordan, Palestine, Tunisia...

- ...male interviewers seem to evoke more agreement when asking if men should have the final say in family decisions ($\chi^2=61.6$, $p<0.01$) – all AB IV countries, except Tunisia
It seems to be more likely that respondents disagree with the statement "women should be able to become prime minister/president" when the female interviewer wears a headscarf ($\chi^2 = 23.4, p<0.01$) – esp. Jordan, Palestine, Tunisia.

Respectively, wearing a hijab appears to evoke more agreement with the statement that men should have the final say in family decisions ($\chi^2 = 8.5, p<0.01$) – all AB IV countries, except Tunisia.
Respondents seem to be more inclined to disagree with the statement that women should be able to become prime minister/president when a third person is present ($\chi^2 = 5.0, p<0.05$) – esp. Algeria, Jordan, Morocco

But, presence of a third person seems to lead to more agreement with the statement that men should have the final say in family decisions ($\chi^2 = 6.3, p<0.01$) – esp. Algeria, Morocco, Tunisia
Measurement: Implications

- Design of a more extensive interviewer questionnaire by the end of the survey interview
- Request for provision of core information about interviewers (ID, name, gender, etc.) from the field agencies
- Interviewer questionnaire capturing similar information (sealed envelope approach) at interviewer training
- Switch to Computer Assisted Personal Interviews in (almost) all countries
- Request for more detailed field reports detailing fieldwork procedures plus relevant records
Is the Arab survey world just a few decades behind?

- Arab survey world is in transition
- Benstead (2018) proposes an extensive agenda for methodological research on the Arab world, but are we ready to walk the walk...yet
- The AB has started collecting relevant data to empirically and systematically investigate representation and measurement error
- The AB V will collect most of these data, so stay tuned...
Contact Information

Kathrin Thomas: kathrint@princeton.edu

Michael Robbins: mdr7@princeton.edu

www.arabbarometer.org
@arabbarometer
@ArabBarometer