Surveying the Arab World: Methodological Challenges and (some) Solutions

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

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

25 August 2018
Where to begin...

"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 [...]. 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)
Just a few decades behind?

- Beginning of survey research and public opinion polling in 1980s, but *increased* interest since the Arab Spring 2011 (Benstead, 2018)
- Limited competitive field organisations and, thus, also limited capacities
- Special political situations may make it more challenging to implement methodological improvements
Holistic framework to study error at every stage of the survey process

Typically, distinguished representation from measurement error

To date, most methodological studies focus on a few aspects, no universal model to study all aspects simultaneously available
The Total Survey Error

MEASUREMENT
- Construct
- Measurement
- Response
- Edited Data

REPRESENTATION
- Population
- Sampling Frame
- Sample
- Respondents

Errors:
- Validity
- Measurement Error
- Processing Error
- Coverage Error
- Sampling Error
- Nonresponse Error

Survey Statistic
The Total Survey Error

**MEASUREMENT**
- Construct
- Measurement
- Response
- Edited Data

**REPRESENTATION**
- Population
- Sampling Frame
- Sample
- Respondents

**Errors**
- Validity
- Measurement Error
- Processing Error
- Fabrication
- New/different kinds of I & R Effects
- Coverage Error
- Sampling Error
- Nonresponse Error
- Outdated Censuses; mobile populations
- Absence of Sampling Frames
- Alternative Sampling Strategies
- Fabrication

**Survey Statistic**
Data and Methods

- Arab Barometer III and IV
- Methodological challenges: Interviewer gender, interviewer religiosity, third party presence
- Representation: Coverage, sampling error, and unit nonresponse; contact attempts; suspected data fabrication
- Measurement: Item non-response and response bias due to interviewer and fieldwork effects
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
- Result: Almost 9 out of 10 interviews in Jordan were conducted by female interviewers
Interviewer Gender

AB IV

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

Source: Arab Barometer IV
Approx. three quarters of the female interviewers wore a hijab during the survey interview (77.0 percent)

<table>
<thead>
<tr>
<th>Country</th>
<th>Percentage</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>
Almost every other interview (!) was attended by a third person (44.6 percent)

Source: Arab Barometer IV
Things to consider...

- We need to be extremely cautious when investigating gender-of-interviewer effects question thinking about the fe/male ratio of interviewers
- We need to be aware that religiosity is often coded for female, but not for male interviewers
- Third party presence may be problematic, but rarely coded or considered
Little empirical evidence about coverage, sampling error, and unit nonresponse

- Coverage tricky, esp. in less stable countries, e.g., Iraq only "more than 66% of the population" covered in AB III
- Extent of sampling error hard to calculate due to lack of relevant information
- Response rates largely AAPOR standard, but not for all countries; non-collection/non-provision of unit nonresponse files, e.g., Palestine 91% response rate (response/response+refusal) in AB IV
Example Tunisia

- Up to 8 contact attempts before Unit nonresponse is logged
- On average, interviewers need 2.8 contact attempts to conduct an interview
- It seems that female interviewers need approximately one fewer contact attempt to interview ($t=2.7; p<0.01$)
Representation: Suspected Data Fabrication
Example 1

Source: Arab Barometer III
Source: Arab Barometer IV
Representation: Suspected Data Fabrication Example 2

Source: Arab Barometer III

Source: Arab Barometer IV
Representation: Suspected Data Fabrication

<table>
<thead>
<tr>
<th>Percent Match</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>
Example Tunisia

- Timestamps recorded: Start/end date and special module stamps
- Shortest interview just over 20 minutes; longest interview just over 80 minutes; average interview length just over 40 minutes
- Interviews conducted by female interviewers tend to be approx. 4 minutes longer (t=-6.4; p<0.01)
Agreement: "Despite negative US foreign policies, most ordinary Americans are good people." (Binary)

Item nonresponse, i.e. proportion of don’t know-and refused-answers, is 16.7 percent on average, but substantive crosscountry variation.

Source: Arab Barometer IV
Interviewers conducted by male interviewers appear to have higher item nonresponse ($\chi^2=26.3; p<0.01$)
- Effect prevalent in all AB IV countries, but Algeria and Lebanon

Item nonresponse seems to be higher when third persons are present ($\chi^2=17.3; p<0.01$)
- This seems to be driven by significant tests in Algeria, Morocco, and Tunisia
Agreement: "A woman can become President or Prime Minister of a Muslim country." (4-point scale)

Agreement higher when female interviewer conducted the interview ($\chi^2=66.3$, $p<0.01$)

But, less agreement when a female interviewer with a hijab conducts the interview ($\chi^2=33.9$, $p<0.01$)

In addition, hesitation to agree when a third person attends the interview ($\chi^2=6.4$, $p<0.05$)
Agreement: "Husbands should have final say in all decisions concerning the family." (4-point scale)

Fewer respondents in interviews conducted by female interviewers agree ($\chi^2=75.3, p<0.01$)

But, more agreement when the female interviewer wears a hijab ($\chi^2=21.1, p<0.01$)

In addition, also more agreement when a third person is present at the interview ($\chi^2=9.2, p<0.01$)
Conclusions

- Systematic methodological research is required to further improve the overall survey data quality.
- Esp. the representation side requires more attention.
- Interviewer and fieldwork effects cannot be neglected.
Implications

**Representation**
- Detailed records of sampling procedures by contract
- Unit nonresponse files by contract
- AAPOR Response Rates by contract
- Live data checking tool

**Measurement**
- Interviewer questionnaire at briefing
- As well as after each interview
- Interviewer details by field organisation
- Computer Assisted Personal Interviews
Benstead (2018) proposes an extensive agenda for methodological research on the Arab world, but we are not quite ready yet.

Arab survey world is in transition, so we can achieve more than ever before.

The Arab Barometer has already started collecting relevant data to systematically and empirically test issues of data quality.

The fifth wave collects extensive data to address methodological challenges. Stay tuned!
Contact Information

Kathrin Thomas:
kathrint@princeton.edu

Michael Robbins:
mdr7@princeton.edu

www.arabbarometer.org
@arabbarometer
@ArabBarometer