Of men and measures. Counting and estimation in comparative survey research.

Agata Grabowska PhD
3MC Berlin 25-28.06.2008
Surrounded by numbers

We all deal with numbers in our everyday life. Dates, hours, id numbers, frequency of events, codes and passwords…

In many comparative survey studies we ask respondents to tell us i.e.

1. How many cups of coffee do you have on a regular weekday?
2. How much time do you spend watching TV every night?
3. What is your annual income?
4. How often do you buy groceries?
5. How much do you spend on gas weekly?
6. What was the amount of alimonies paid you since Jan 1, 2008?

…and so on

(sources: questionnaires for ESS, EQLS, IPOS OMNI, SIRS)
Do we think in numbers?

Most people are bad with all kind of numerical estimates and they either over- or underestimate time, spendings or occurrence of events. Moreover, there seem to be significant cultural differences?

Sceptical? Try this?

- How long is one meter in Poland, Ukraine and Lithuania?
- How many cigarretes do Poles and Americans smoke?
- How far is not-so-far in Poland and Greece?

There is an impressive body of research on frequency of event/behavior estimation [see: Linton:1992, Wagenaar: 1986, White 1992] which proves that:

- increased frequency of event/behavior occurrence
- better semantic memory
- decreased accuracy and ability to differentiate occurrence
Retrieval strategies

1. Counting
2. Estimation
3. Anchoring

Choice of a particular strategy depends on a number of factors:

- frequency of events/behaviors
- time elapsed/time frame
- regularity
- similarity
- question design (verb, syntax)
Accuracy and errors

Errors result mostly from:

- cognitive burden
- effort input
- time period
- event/behavior characteristics
- illusion of stability

Counting

- omission
- comission

Estimation

- illusion of stability
Research procedure

To see whether people do in fact make errors in estimating/counting we need to use external validation.

- self generated data
- observation
- external data
Experiment I
Experimental design

**Goal:** compare accuracy of counting vs percentage of Polish and Italian speakers estimating duration (110 subjects)

**What:** duration of morning rounds over one week period

**Where:** Milan (Ospedale di Limbiate) and Warsaw (Szpital MSWiA)

**When:** March 2008

**Procedure:** Team of 8 interviewers observation journals and the conducted cognitive interviews (supervision for Poland Agata Grabowska, for Italy – Simone Malto)

\[ N \text{ total} = 110 \]

\[ N(p) = 42, \ N(i) = 58 \]
Experimental design

After a week had passed half of the doctors where asked to count the time they spent doing morning rounds that week and half to give a percentage of their time spent doing so.

We used thinkalouds to monitor the retrieval process.

“So it must have been like 40 minutes at the most, because I ran form the ER at around 7.15 and did not even have time for coffee and we were done by 8 because my mother called then and I was already in my office. Yeah, roughly 40 minutes on Thursday.” (Polish respondent, male, 42)

Additionally we asked doctors to assess how difficult the task of retrieval was:

1 very easy
2 rather easy
3 neither easy nor difficult
4 rather difficult
5 very difficult
### Results

<table>
<thead>
<tr>
<th></th>
<th>Polish</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>counting</td>
<td>4 h 05 min (3.1)</td>
<td>6 h 05 min (4.3)</td>
</tr>
<tr>
<td>percentage</td>
<td>7 h 45 9 min (1.7)</td>
<td>8 h 45 min (1.6)</td>
</tr>
<tr>
<td>real time</td>
<td>3 h 45 min</td>
<td>4 h 25 min</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Polish</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>error (c)</td>
<td>20 min</td>
<td>1 h 40 minutes</td>
</tr>
<tr>
<td>errror (p)</td>
<td>3 h</td>
<td>4 h 25 min</td>
</tr>
</tbody>
</table>

Verbs used: active voice 1 person singular (*dedicar tiempo a* and *poświęcać czas na*)
Experiment II
Experimental design

**Goal:** compare accuracy of duration estimation of Polish and British Tesco shoppers (400 subjects)

**What:** waiting in line to the checkout

**Where:** Tesco supermarkets in Warsaw and Glasgow

**When:** June 12th, 2008
Experimental design

**Procedure:** Team of interviewers observed Tesco shoppers waiting in line and measured the exact time passed. Then they conducted a 3 minute exit interview with these shoppers. Supervision for Poland – Agata Grabowska, for the UK – Christian Cork.

**Sample:** quota sample of 200 male and female adult Tesco shoppers (we controlled age, family status and # items bought).

Additionally we had shoppers assess their shopping experience and the waiting time. For waiting we used a 4 point scale:

1. it was fine
2. not a big deal
3. rather irritating
4. very irritating
### Results

|                      | Polish  
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N= 199)</td>
</tr>
<tr>
<td>estimation</td>
<td>8 min 45</td>
</tr>
<tr>
<td>real time</td>
<td>5 min 10 sec</td>
</tr>
<tr>
<td>error</td>
<td>3 min 35 sec</td>
</tr>
<tr>
<td>irritation</td>
<td>4.2</td>
</tr>
</tbody>
</table>

- older shoppers were less irritated and accurate
- individual shoppers were more accurate than family shoppers
- shoppers buying less items were less accurate (Poland)
- frequent Tesco shoppers were more accurate
- women were less irritated (2.9 vs 4.1) more accurate than men
Results

The longer people waited the less accurate they were.

N= 199 (Poland)
Conclusion

People generally tend to give inaccurate estimations of duration and these are heavily influenced by retrieval strategy, task difficulty, nature of event and emotions associated with it.

Errors of over- and under estimation are put into national database

These errors may be additionally increased by cultural factors, which do not rarely control.

We work on international data and publish comparative reports.
Thank you

Agata Grabowsk, PhD
agata.grabowska@4prm.com