Coding Verbal Behaviors to Measure Data Quality in Interviews Collected with Conventional Questionnaires and Event History Calendar

Mario Callegaro
*Program in Survey Research and Methodology (SRAM)*
University of Nebraska, Lincoln

Workshop on Behavioural Coding, 16 February 2007
Wivenhoe House, University of Essex

---

Conventional questionnaire

- Standardized interviewing
- The first behavioral coding scheme was used to study how interviewers departed from standardized interviewing
- In Conventional Questionnaires interviewers are not allowed to change the order of the questions
Interviewing with EHC

- The interviewer guides the respondent in answering queries for each time line
- The process uses information and dates for each completed domain to help the respondent correctly place other events in the appropriate time frame
- Interviewers use scripts more than standardized questions
- The unit of analysis can be of years, months, and thirds of a month

The interview process

- The entire process of compiling the calendar focuses, by its nature, on
  - Coherence
  - Consistency
  - Sequential nature
  - Attempts to correct for missing data
- The time line also highlights missing data in a more prominent way than a conventional questionnaire
EHC from a memory point of view

- Provides structure of timelines and domain themes that reflect the structure of autobiographical memory
- Facilitates the use of three memory retrieval mechanisms to more completely and accurately reconstruct the past
  - Top-down associations
  - Sequential associations
  - Parallel associations
- Allows the use of a more naturalistic narrative approach to remembering
- Encourages motivation to remember as the inherent cueing mechanisms lends retrieval to be more productive

The research questions

- Are retrieval cues more prevalent in EHC interviews?
- Retrieval cues can be parallel, sequential, and top-down
- Assess the role of conversational processes more prevalent in EHC interviews
- Examine behaviors that indicate negotiation of uncertainty with regards to question and response meaning
- Examine behaviors that indicate rapport in I/R interpersonal relationship
Final goal

- Gain insights into both retrieval and conversational processes that could potentially affect data quality
- What are the interviewer/respondent behaviors that can improve the quality of retrospective reports?

Study’s methodology

- A subset of Panel Study of Income Dynamics respondents and 20 interviewers were randomly assigned to the EHC or the CQ condition
- EHC cond. N=309, 84.4% AAPOR COOP1
- CQ cond. N=307, 84.1% AAPOR COOP1
- 2-year reference period
- 7 Domains:
  - Places of residence, household composition, employment, earned income, unemployment and not in the labor force, time away from work, entitlements
PSID EHC, 1998 (Belli et al., 2001)

### Residential

### Household composition

### Employment

### AFDC

<table>
<thead>
<tr>
<th>Week 1996</th>
<th>Year 1996</th>
<th>Season 1996</th>
<th>Landmark Events</th>
<th>Street City State</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>February</td>
<td>March</td>
<td>April</td>
<td>May</td>
</tr>
<tr>
<td>Neo School</td>
<td>20th</td>
<td>4th</td>
<td>Independence</td>
<td>2nd</td>
</tr>
<tr>
<td>[3 1/4 years ago]</td>
<td>[2 years ago]</td>
<td>[1 3/4 years ago]</td>
<td>[1 1/2 years ago]</td>
<td></td>
</tr>
<tr>
<td>Did the respondent move because...</td>
<td>Did the respondent move because...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A was on vacation or took time off in 1996 Y</td>
<td>A time vacation or took time off in 1997 Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If yes, from...</td>
<td>If yes, from...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A was temporarily laid off in 1996 Y</td>
<td>A was temporarily laid off in 1997 Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If yes, for how long...</td>
<td>If yes, for how long...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A was sick in 1996 Y</td>
<td>A was sick in 1997 Y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Someone else was sick in 1996 Y</td>
<td>Someone else was sick in 1997 Y</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**1996 Time Away from Work:**
- Did the respondent move because...
- Did the respondent move because...

**1997 Time Away from Work:**
- Did the respondent move because...
- Did the respondent move because...

**1996 Earnings Status:**
- Imputed earnings for 82 week period in 1996 Y
- Imputed earnings for 82 week period in 1997 Y

**1997 Earnings Status:**
- Imputed earnings for 82 week period in 1996 Y
- Imputed earnings for 82 week period in 1997 Y
Verbal behavior coding

- 95% of interviews taped and transcribed
- Extensive coder training and weekly meetings with the principal investigators
- Because of refusal, inaudible tapes and attrition of the coders
  - 70.5% of EHC interviews were coded (218)
  - 64.1% of CQ interviews were coded (197)
- 38 interviews (9.2%) blindly double coded to assess intercoder reliability

Initial coding scheme

- First coding scheme for EHC
  - No clearly defined question/answer sequences
- Staring point:
  - CQ behavioral coding schemes that identify cognitive problems (Fowler & Cannell, 1996; Oksenberg, Cannell & Kalton, 1991)
  - CQ behavioral coding schemes that indicate response accuracy, retrieval and conversational behaviors (Belli, Lepkowski, & Kabeto, 1999; Houtkop-Steenstra, 2000*)
Development of the coding scheme

- Coding team:
  - Principal investigator (prof. Belli), 1 EHC and 1 CQ professional interviewer, 2 senior undergraduates with no prior experience in coding or interviewing

- First phase: finalizing coding scheme
  - 10 weeks, the 4 coders working 6-7 hours a week
  - Weekly meetings analyzing 8 transcripts (5 EHC, 3 CQ)
  - Selective coding training (Ongena, 2002)

- Second phase: coder training and developing inter-coder reliability
  - 12 weeks, the 4 coders working with 12 transcripts (8 EHC, 4 CQ)

Production coding

- 56 verbal behaviors
- Domain as unit of analysis
- Reliability analysis of the 38 blindly double coded interviews excluded 16 behaviors
- Some behaviors were excluded for they very low occurrence (< 10)
- Analysis done on 40 verbal behaviors
- Only the two undergraduates served as coders
Coding scheme I

- Interviewer’s behaviors classes:
  - Retrieval questions/probes
    - Parallel retrieval
    - Sequential retrieval
    - Top-down retrieval
    - Miscellaneous retrieval
  - Calendar year questions/probes
  - Uncertainty behaviors
  - Problem with standardizations
  - Response to perceived respondent cognitive difficulty
  - Feedback
  - Rapport

Coding scheme II

- Respondent’s behaviors classes
  - Retrieval strategies
    - Parallel
    - Sequential
  - Response to an interviewer uncertainty
  - Cognitive difficulty
  - Miscellaneous behaviors
  - Rapport
  - Response to interviewer narrowing probe
I’wer Retrieval probes (1)

- Parallel
  - interviewer referring to a contemporaneous event from respondent’s past
  - Personal, holiday, or historical
  - E.g., “when you got married”; “do you remember if that was before or after John Kennedy was shot?”

I’wer Retrieval probes (2)

- Sequential
  - Duration – seeking how long, or how much time
    - E.g., “how long did you live at the next residence?”
  - Timing – seeking to know when a spell (period) began or ended
    - E.g., “do you remember what month it was?”
  - Unspecified sequential
    - E.g., “where did you go after?”
I’wer Retrieval probes (3)

- Data Elements
  - Seeking specific required information within a spell
    - E.g. “May I have that address?”

I’wer Conversational Probes

- Verification
  - Interviewer is making certain the respondent’s prior answer(s) are correct
  - E.g., “I have that you lived in the same street again, at a different address, is that true?”

- Seeking clarification
  - E.g., “Were you like on a vacation there?”

- Directive probing
  - Any probe that can possibly bias an answer
I’wer Conversational Verbal Behav.

- **Feedback**
  - Task related
    - E.g., “Give me a second to get to the next spot, sir”
  - Acceptable and unacceptable
    - E.g., unacceptable, “Boy, that must have been awful”
  - Digressions
    - E.g., “oh, hang on to those, that will be invaluable later on for your family”
  - Laughter

Respondent Retrieval Strategies

- **Coded when unprompted**
  - Parallel (personal, holiday, historic)
    - E.g., “I lived at uhm, for 6 months after I got married in Michigan”
  - Sequential
    - Duration (“only for about 6 months), Timing (“when I was 18), and undifferentiated sequential (“I went to work in a doctor’s office”)
Other respondent behaviors

- Agreements and disagreements to verifications and directives
- Cognitive difficulty
  - Requests for clarification (e.g., “now you said technical I went to x-ray training”)
  - Qualified responses (e.g., “I guess”)
  - Does not meet
  - Don’t know
- Conversational
  - Explanations, digressions, and laughter

Excerpted Example: Labor History

I: And how long did you stay there please?  Duration probe

R: Until, um, October of ’92.  Timing Response

I: Okay. And then, um, in October of ’92, did you take another job?

R: I took another job of—but it lasted for, like, a month and then I went to work some place else — your interested in— Sequential response

I: Not that month one then, we’ll take the next  Clarification
Excerpted Example (continued)

Sequential response
R: Okay, um. The next job was at, um, let's see. Trying to think of the name of the place, they changed their name. Um. They used to be called employer10?

Data element

Exhibition
I: Employer10?

Verification probe
R: Yes.

Verification agreement
Timed probe
I: And when did you start working for them? In '92?
R: Um. Yes, December of '92 until May of '93.

Directive agreement

Timing probe

Timing dual

Excerpted Example (continued)

Sequential probe
I: Alright. December of '92 until May of '93. And then in May of '93 did you…?

Data element response
R: Um, I went to work in state3 for employer11, but I'm trying to think how long I worked there. It wasn't—it was—it was until—let's see. I came to state3 in August—I must have gone for a job—oh, I guess I started in January of '94, and I only worked for 6 weeks, and then I went on disability.

Exhibition

Duration response
Parallel response

Timing response
Final coding scheme

- 41 of the original 56 verbal behaviors met criterion of intercoder reliability (r ≥ .40)
- Factor analysis (principal components) identified 4 latent factors (eigenvalues > 2)
  - Retrieval Cues
  - Detailed Interviewing
  - Cognitive Difficulty
  - Conversational Rapport
- The factors are composed of interviewers and respondents' behaviors

Lessons learned

- The professional interviewers were having difficulty in assigning codes that were as reliable as the coders. Accordingly, only the 2 undergraduates served as coders.
- Coding was done on the transcripts only.
- Some years later bad transcriptions (Poland, 2003) were found and 36 CQ transcriptions could not be used (Callegaro et al, 2007).
- Fortunately the 36 CQ interviews did not change the main study outcome.
Bad transcripts

- After investigation, it appeared than one transcriber was the problem
- To save time CQ transcribers used the questionnaire as a draft
- One transcriber thought he/she had to leave out what was not strictly question/answer dialogue
- Monitor transcribers and make sure to record “who transcribed what” in order to go back if necessary

Some Results (Belli et al, 2004)

- There is not a significantly greater degree of interviewer variation on data quality in the EHC method
- Better data quality was associated with a higher prevalence of retrieval cues, a greater degree of response openness, and lower levels of cognitive difficulty and rapport
- The association of data quality and verbal behavior also interacted with method: retrieval cues and cognitive difficulty were directly associated with EHC response quality and indirectly associated with Q-list quality; rapport behaviors had a more detrimental effect on Q-list data quality
Some results (Callegaro et al, 2007)

- Overall African American respondents more frequently showed behaviors indicating cognitive difficulty
- At the same time there was no difference in the level of cognitive difficulty between CQ and EHC interviewing methods, regardless of race
- Conversational rapport was more frequent in the EHC condition especially for the combination of African American interviewers and African American respondents

Future direction

- New coding scheme now applied
- Promising results in intercoder reliability
- Higher intercoder reliability in interviewer behaviors than in respondent behaviors