

The Survey Response Process from a Cognitive Viewpoint

Roger Tourangeau, Westat

Seminar on "Improving the Quality of Data Collection in Large Scale Assessments", 11 May 2017, Paris, France.

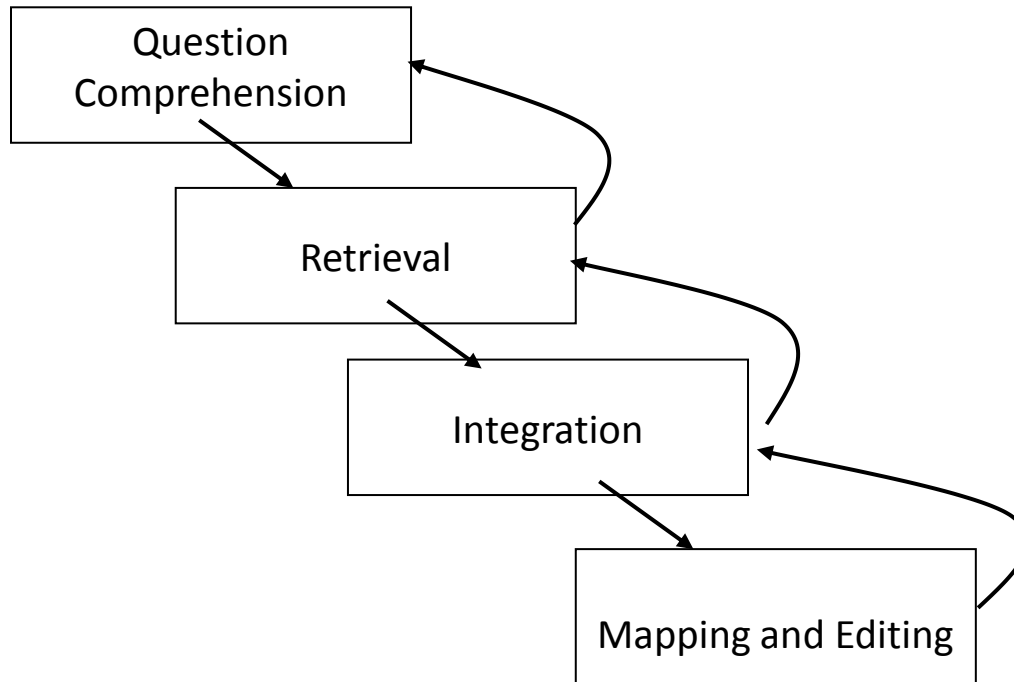
Outline for Today's Talk

- Components of the response process--the ESCRIME model
 - Encoding
 - Storage
 - Comprehension
 - Retrieval
 - Integration/Estimation
 - Mapping
 - Editing
- Understanding each component suggests ways to reduce errors
- Alternatives to the ESCRIME model

Measurement Error

- How does it arise?
- Main source appears to be R giving incorrect answer (How do we know this?)
- Other possibilities include:
 - Errors in recording answers
 - Transcription errors
 - Errors in coding
 - Data entry errors
- Difficult to separate q'aire and R as sources of error

Simple Model of the Response Process



- Sequential, although R can backtrack
- Models performance of ideal R: Rs may take shortcuts: satisficing, acquiescence

Defining Each Component

- Encoding/acquisition: Noticing and interpreting some aspect of an experience; how information is taken in initially
- Storage: Retention of information in LTM; formation of episodic memory appears to require a period lasting several days (“consolidation”) and to involve specific brain structures
- Comprehension: Understanding the task and the question
- Retrieval: Recalling information from memory, bringing info. into consciousness, explicit memory
- Integration/Estimation: Combining, supplementing, extrapolating from information that has been retrieved
- Mapping and Editing (Reporting): Formatting the judgment/estimate; altering it; conveying an answer via an overt response

Defining Each Component — II

- Each component a package of subprocesses; e.g., comprehension involves processes at the word level, sentence level, and utterance level

One versus Two Tracks

- Cannell, Miller, & Oksenberg model good example of two-track model (one process for good answers, a second process for bad answers)
- Other recent examples
 - Strack & Martin (1987)
 - Krosnick satisficing model (1991): Satisficing (weak and strong) vs. optimizing
- Drawbacks to two-track models: Many paths to an answer
 - Different components skipped or shortchanged
 - Backtracking possible
 - Cognitive toolkit —response process for different questions constructed from a set of common processes
- Still, two-track models popular in psychology (e.g., heuristic vs. systematic processes in persuasion, judgment; Kahneman's *Thinking Fast and Slow*)

Errors a Byproduct of Each Component

- Failure to notice key information or to encode it correctly
- New information woven into representation of an event while it's in storage
- Question misunderstood
- Relevant information forgotten
- Information incorrectly reconstructed or poorly estimated
- Answer is a deliberate misreport

Encoding

- Some events never noticed at all—duration, drama, distinctiveness relate to likelihood of encoding and depth of encoding
- Encoding specificity principle: Mismatch between retrieval cue and encoding in memory produces retrieval failure: “How often do you do light or moderate activities for at least 10 minutes that cause only light sweating or a slight to moderate increase in breathing or heart rate?”
- Related idea: Context-dependent learning (mood-dependent memory as a special case)

Failure to Encode

- Lee et al. (1999) demonstrated that parents remember little about children's vaccinations
 - compared parents' reports to medical records
 - report up-to-date for 80% children who are not up to date
 - recall is poor immediately after injection and no worse after 10 weeks
 - recognition is no better than recall
 - childhood injections are frequent, not particularly distinctive, occur in batches and involve esoteric terms
 - Parents may simply have not encoded enough to accurately recall and report the events

Storage

- Even if accurate information gets into LTM, it can be distorted over time
- Retrieval related to memorability (=accessibility or strength); this in turn related to rehearsal (=thinking/talking about the event)
- Post-event information may be incorporated into memory representation
 - Later events, information may overwrite earlier material
 - A generic memory may form, making it difficult to remember specific of individual events
- Neisser and Harsch on Challenger disaster—flashbulb memories are vivid and detailed, but not necessarily all that accurate
- Source monitoring: Marcia Johnson and colleagues: Where did information come from—direct experience, imagination, reading, second-hand report?
 - Plausibility
 - Vividness, perceptual detail
 - Strictness of criterion

Comprehension Problems

- Tourangeau, Rips, and Rasinski (2000) identify seven types of comprehension problems
 1. Ambiguity and conceptual variability
 2. Excessive complexity
 3. Vague concepts
 4. Faulty presupposition
 5. Vague quantifiers
 6. Unfamiliar terms
 7. False inferences

Two Major Types of Problem

(1) Misalignment

People applying a definition that doesn't agree with "official" definition behind question (residence, smoking a cigarette)

(2) Poor fit

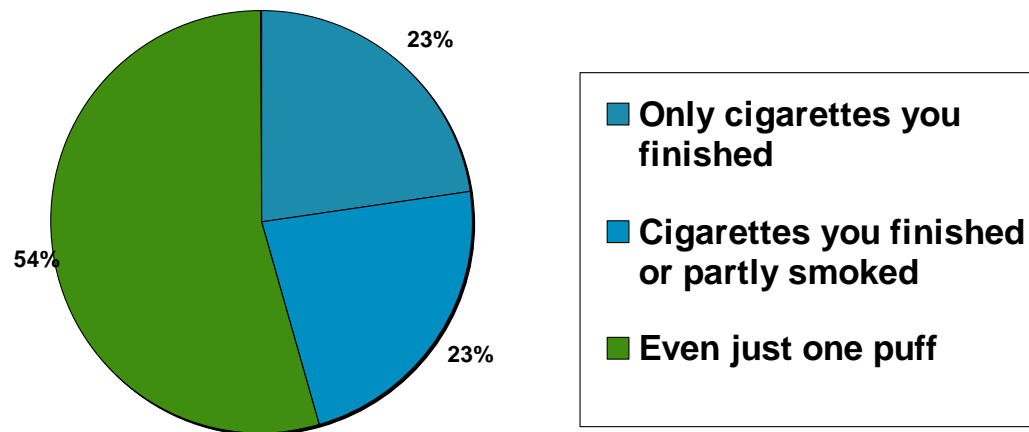
I: Last week, did you have more than one job including part-time, evening or weekend work?

R: Um... I babysit for two families. Is that one job or two?

- Unclear how to apply "more than one job" to ones circumstances
- Every concept has a gradient of instances that clearly fit, clearly don't fit, and are the border line

Comprehension Problems: Conceptual Variability (Misalignment)

- Words have many meanings (senses)
- Suessbrick, Schober & Conrad (2000) administered CPS tobacco use supplement followed by post-test about interpretation:
 - most frequent interpretation held by only 53.8%



Do you think that children suffer any ill effects from watching TV with violence in them, other than ordinary Westerns? By children I mean people under 14, by ill effects I mean increased aggression at school or at home, increased nightmares, inability to concentrate on routine chores, and so on. By violence I mean graphic depictions of individuals inflicting physical injuries on themselves or others, depictions of individuals wantonly damaging property or possessions, abusive behaviors or language to others, and so on.

Forgetting and Other Memory Problems

- Four forms of memory failure
 1. Mismatches between terms in question and terms used to encode events initially
 2. Distortions in the representation over time
 3. Retrieval failure
 4. Reconstruction errors

Memory Problems: Mismatch

How often do you do light or moderate activities for at least 10 minutes that cause only light sweating or a slight to moderate increase in breathing or heart rate? (NHIS)

- If R did not encode vacuum cleaning or walking to work as “light to moderate activity,” might not come to mind

Memory Problems: Distortions Over Time

- Source amnesia
 - Hard to distinguish whether information was actually experienced or added through retelling or thinking about event afterward
 - Even inferred aspects of events are hard distinguish from actual aspects of events:
 - Experimental participants watched film of traffic incident
 - “How fast was the car going when it went through the yield sign”
 - Led to reports of yield sign in original traffic event on a subsequent memory test even when one was not present (Loftus, 1979)

Memory Problems: Retrieval Failure

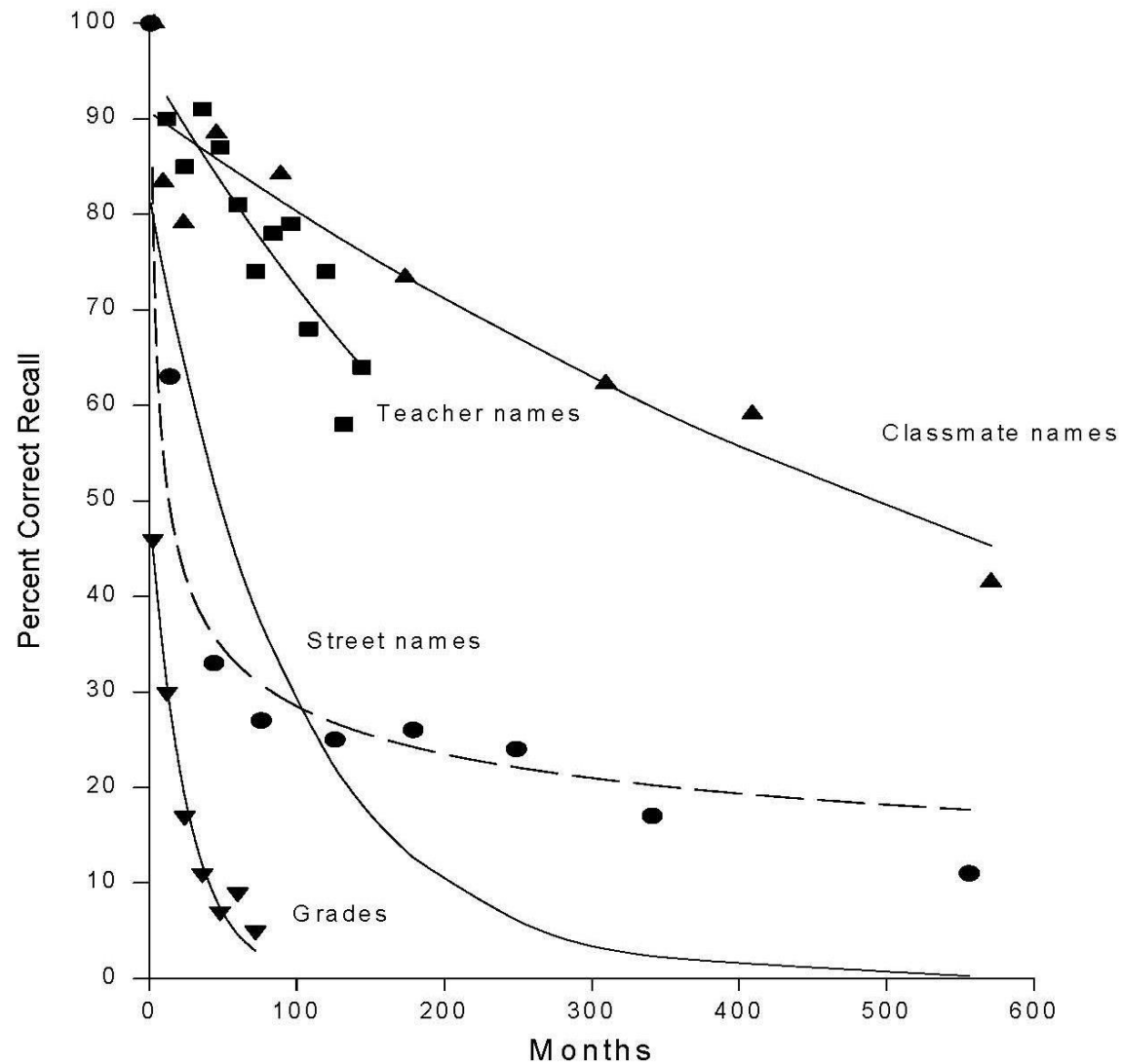
- Interference

- The longer the time period in question (e.g. 1 year vs. 1 month) the more likely other similar events will have occurred
- Hard to distinguish details of one event from others
- Tend to blend into single generic memory

- Decay

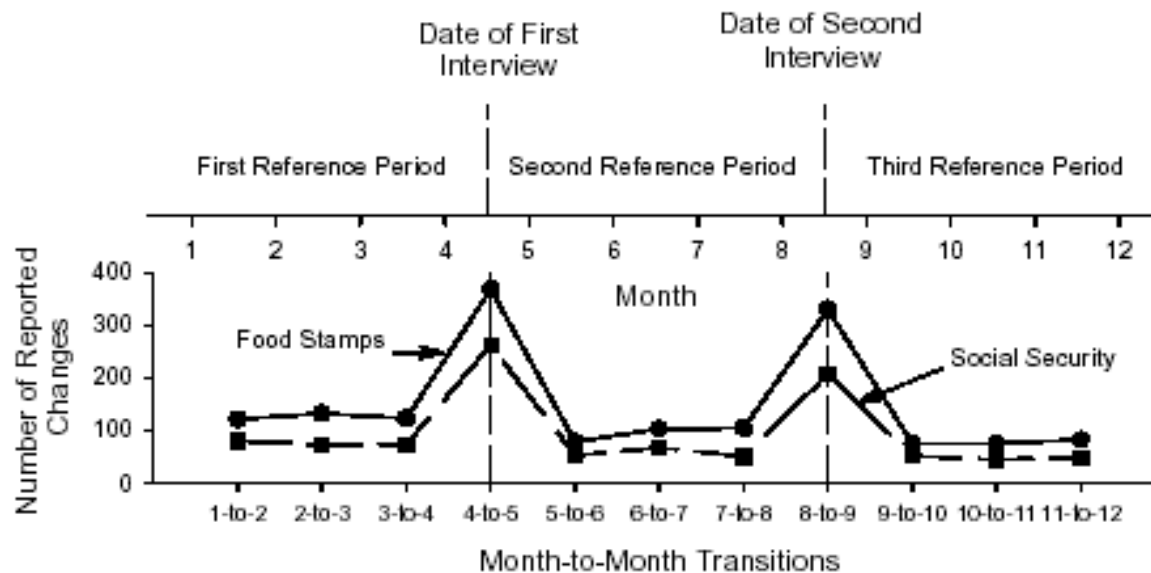
- The more time that has passed since events occurred, the weaker the memory
- Forgetting most rapid in period immediately after event experienced
- Forgetting continues after as many as 50 years(!)

Recall Accuracy for Types of Personal Information



Reconstruction Errors

- The seam effect: What is it?
- Why does it happen?
- Forgetting plus reconstruction
 - Retrospective bias
 - Constant wave response—accounts for seam effect



- Month to month change in Food Stamp and Social Security usage in SIPP

Improving Recall

- Almost all approaches involve providing more or better cues

... tell me if any of these happened to you in the last 6 months, that is since [DATE].

Was something belonging to YOU stolen, such as –

- (a) Things that you carry, like luggage, a wallet, purse, briefcase, book -
- (b) Clothing, jewelry, or cellphone –
- (c) Bicycle or sports equipment -
- (d) Things in your home - like a TV, stereo, or tools –
- (e) Things outside your home such as a garden hose or furniture -
- (f) Things belonging to children in the household -
- (g) Things from a vehicle, such as a package, groceries, camera, or CDs - OR
- (h) Did anyone ATTEMPT to steal anything belonging to you?

Did any incidents of this type happen to you?

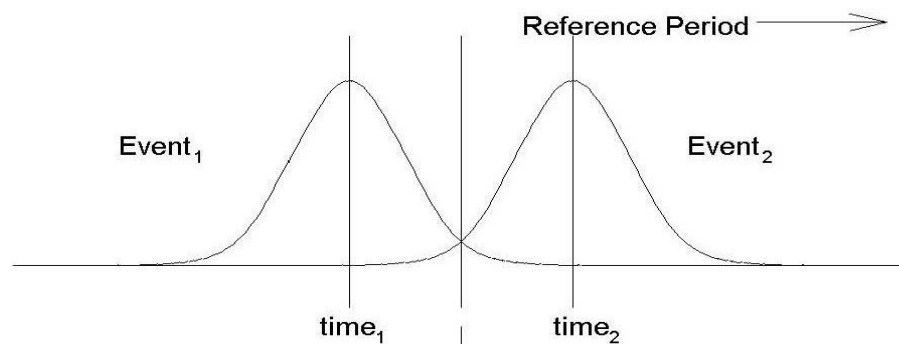
Other Approaches to Improving Recall

- Life events/event history calendar
- Gather rich cues about life events from R; record events and dates on calendars
- Helps Rs retrieve other memories

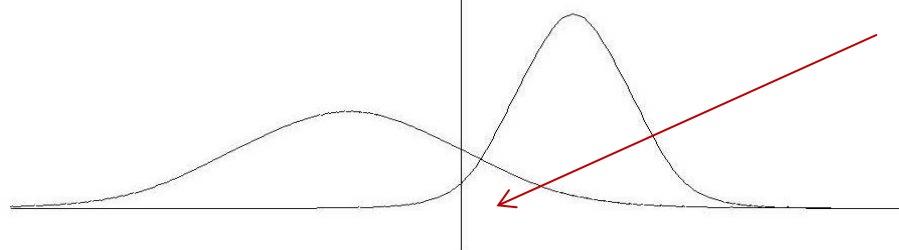
Telescoping Errors

- Variance theories—lose temporal information over time so date has bigger variance with earlier events; e.g., have impression about range of possible dates (Huttenlocher, Hedges, & Prohaska, 1988; Rubin and Baddeley, 1989)
 - Independent evidence that magnitude of data errors increases

a.



b.



More events telescoped into the reference period than telescoped out

Impact of Telescoping

Study	Bounding Procedure	Ratio of Events Reported: Unbounded over Bounded	
Neter & Waksberg (1964)	Prior interview	Expenditures	1.40
		Jobs	1.55
Loftus & Marburger (1983)	Landmark event	Any victimizations	6.15
		Victim of theft	1.51
		Victim of assault	1.52
	Personal landmark	Reported crime	1.22
		Any victimizations	5.50
		Any victimizations	2.00
	New Year's Day	Any victimizations	2.52
		Specific Date	1.32

Impact of Telescoping

Study	Bounding Procedure	Ratio of Events Reported: Unbounded over Bounded	
Sudman, Finn, and Lannom (1984)			
Study 1	Earlier period (previous month) in same interview	Disability days	1.20
		Bed days	1.24
		Visit to medical provider	1.07
		Nights in hospital	1.24
Study 2	Earlier period (previous weekend) in same interview	M-W	
		Mean # snack purchases	1.49
		Mean \$	1.34
		Weekends	
		Mean # snack purchases	1.85
		Mean \$	2.19

Estimation Problems: Behavioral Frequency Questions

Now think about the past 12 months, from [DATE] through today. We want to know how many days you've used any prescription tranquilizer that was not prescribed to you or that you took only for the experience or feeling it caused during the past 12 months. (NSDUH)

- At least three broad strategies, each leading to different type of error
 1. *Recall and count*: underestimation; may also be prone to overestimation due to telescoping
 2. *Rate-based estimation*: overestimation
 3. *Impression-based estimation*: overestimation

What are Estimates/Judgments Based On?

Basis for Answer	Temporal Question	Frequency Question	Attitude Question
Existing answer	Exact date	Tally	Existing judgment
General information	Temporal period	Rate, other generic info.	Values, predispositions
Specific information	Other details "Recall-and-infer"	Specific episodes "Recall-and-count"	Specific beliefs, "considerations" "Recall-and-integrate"
Impression	Impression	Impression	Impression, stereotype

Formatting Problems

- Three general format types:
 1. Open, numerical responses
 2. Closed with ordered response scales
 3. Closed with categorical response options

Formatting Problems:

Open Numerical Responses

- Problems with Open Numerical format

Now thinking about your physical health, which includes illness and injury, for how many days during the past 30 was your physical health not good? (BRFSS)

- May be hard to convert vague impression into number
- Rounded numbers may indicate difficulty with conversion or signal that R is estimating

Formatting Problems: Ordered Response Scales

- Problems with Ordered Response scales

Would you say that in general your health is

1 Excellent

2 Very Good

3 Good

4 Fair

5 Poor

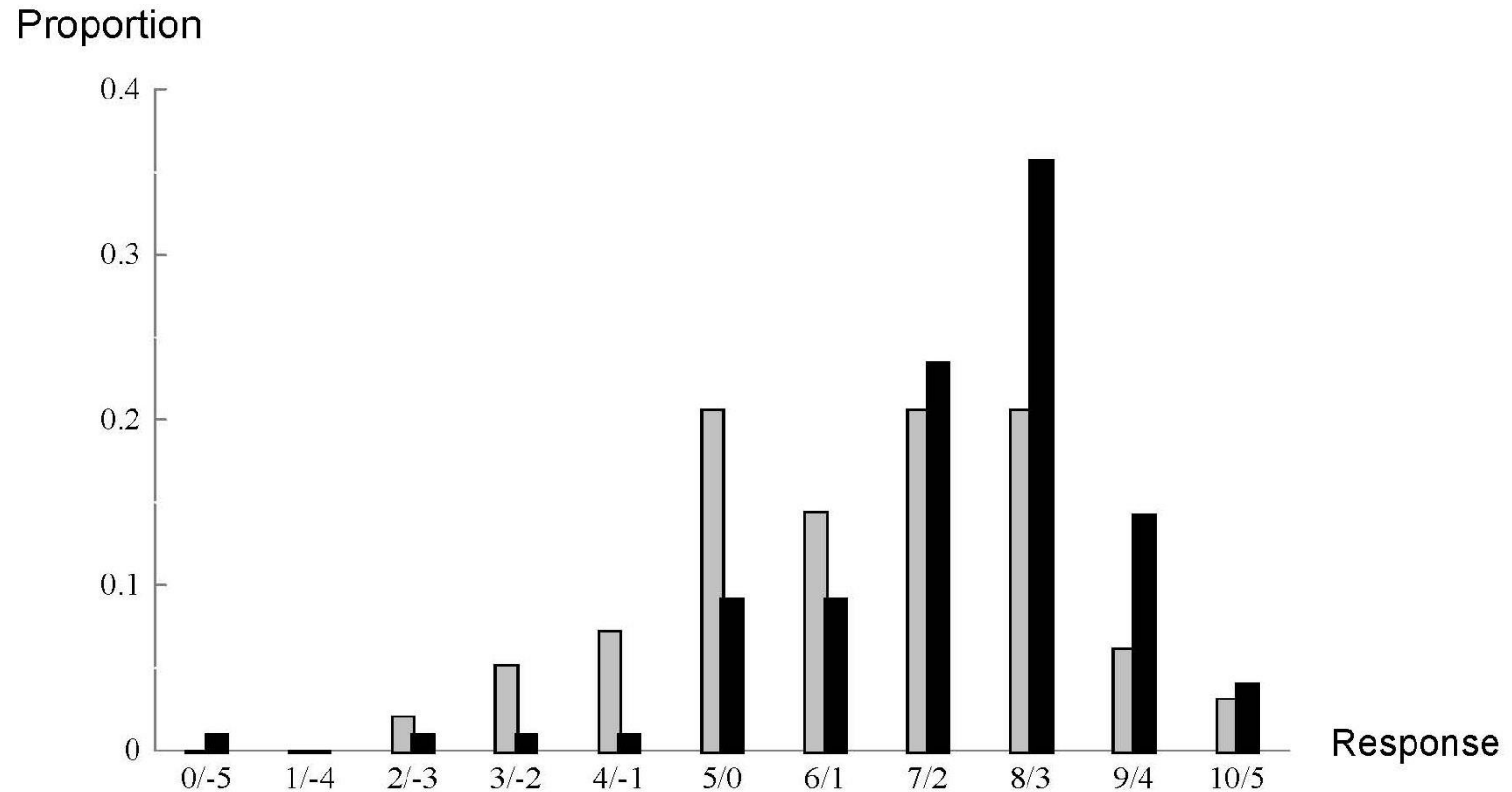
(BRFSS)

- Positivity Bias:

- Rs tend to endorse more positive than negative values
- Schwarz, et al. (1991) suggested more extreme when numerical labels lead to more negative interpretation of low end verbal labels



The Effect of Numerical Labels



The black bars represent the distribution of responses on -5 to +5 scale; the gray bars represent responses on a 0 to 10 scale. (Data from Schwarz et al., 1991.)




Formatting Problems: Unordered Response Scales

- Problems with Unordered Response Scales

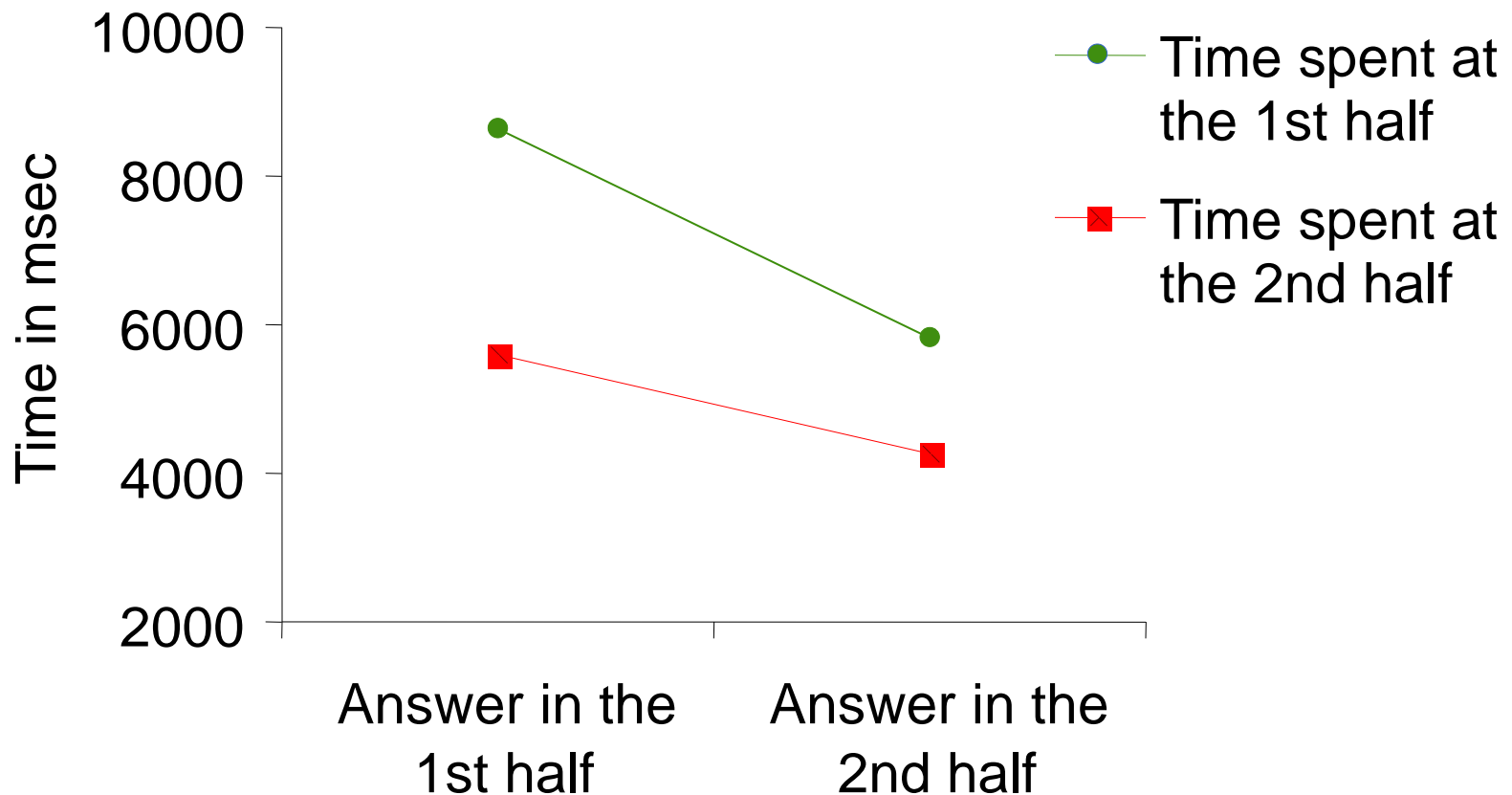
Are you:

- 1 Married
 - 2 Divorced
 - 3 Widowed
 - 4 Separated
 - 5 Never married
 - 6 A member of an unmarried couple
- Recency effect: tendency to endorse last option in list
 - most likely when interviewer reads to respondent
 - Primacy effect: tendency to endorse first option in list
 - most likely when respondent reads to self or when predictable scale being used

Ways to Answer the Question

- Considering all options and choosing the best answer:
web14 
web14_q5.avi
- Selecting the first option, then going through the list and updating the response: web03 
web03_q5.avi
- Reading only part of the list, then selecting the answer:
web01 
web01_q5.avi

Time Spent Looking at Response Options



Note: Time corrected for the time needed to point to an answer (800 msec) and click on it (200 msec); cf. Kieras, 2001; John & Kieras, 1996

Editing

- Deliberate alteration of response (motivated misreporting)
- Social desirability
- Burden reduction (avoiding follow-up questions)

Reporting Errors

- Overreporting (Voting, Church attendance)
- Underreporting (Illicit drug use, abortion, smoking, alcohol consumption, racist attitudes)
- Both overreporting and underreporting (sexual partners)

What Helps Improve Reporting?

- Since the mid-70's have known some basic truths
 - Self-administration helps (addresses concerns about disclosure to an interviewer)
 - Open items better than closed items
- Bogus pipeline
- RRT and ICT both have mixed track record

Conclusions—Editing

- Misreporting gets worse as items get more sensitive
- Misreporting worse among those with something to hide
- Three things help:
 - Self-administration
 - Bogus pipeline
 - Open questions
- One thing hurts: Parental presence
- People trying to avoid embarrassment: Public disclosure of negative information to a stranger
- People lie in surveys for the same reasons they lie in everyday life—to spare themselves embarrassment and to spare other people's feelings

Merci beaucoup!!!!