

The Psychology of Survey Response

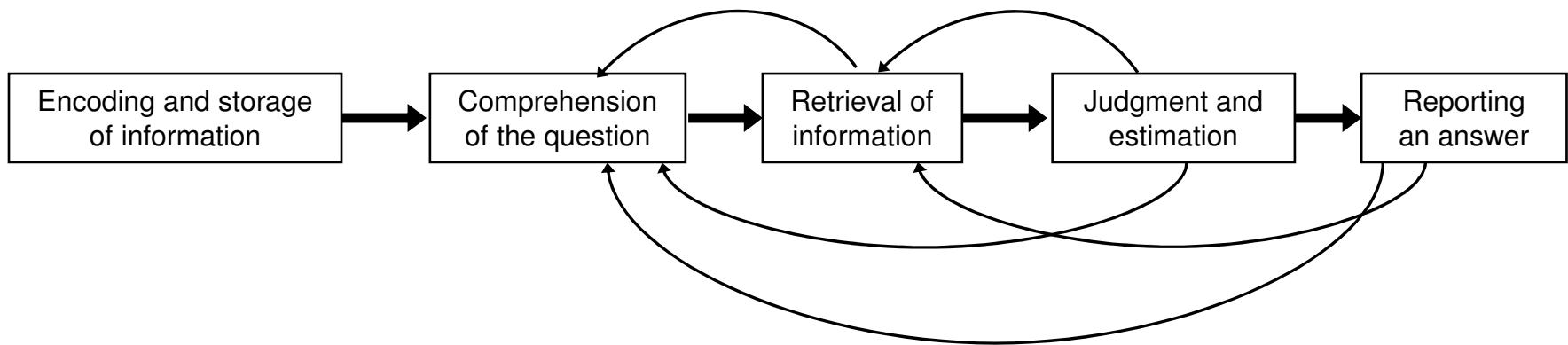
An ASA Webinar
February 9, 2010

Outline for Today's Class

- Components of the response process--the ESCRIME model
 - Encoding } Before Q is asked
 - Storage }
 - Comprehension
 - Retrieval
 - Integration/Estimation } During or after Q
 - Mapping
 - Editing
- Main point: Reporting errors arise because of problems at one or more stages

The Response Process

Simple Model of the Response Process



- Sequential, although R can backtrack
- Models performance of ideal R:
 - Rs may take shortcuts: satisficing, acquiescence
 - Rs may intentionally misreport: social desirability
 - Event(s) may not be recorded (“encoded”) in R’s memory

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)

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 and Storage Issues

Encoding

- Some events never noticed at all—duration, drama, distinctiveness relate to likelihood of encoding and depth of encoding
- Depth of encoding (Count the number of occurrences of the letter “A” vs. summarize the passage)
 - People recall what they encode in the terms in which they originally encoded it
 - Amount may reflect level of elaboration
 - Memory is a network; retrievability reflects number and strength of links
- Encoding specificity principle: Mismatch between retrieval cue and encoding in memory produces retrieval failure (pregnancy vs. abortion in NSFG)
- 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

Lee et al. Results

Immunization Study: Wave 1 Results

Vaccine	Accuracy Measure			
	False Negative	False Positive	Phi	Net Bias
Hep B	51.7 (60)	20.0 (10)	.20 (70)	-41.4 (70)
DTP	41.4 (58)	16.7 (12)	.32* (70)	-31.4 (70)
Polio	33.9 (56)	14.3 (14)	.42* (70)	-24.3 (70)
Hib	86.5 (52)	0.0 (18)	.20 (70)	-64.3 (70)
MMR	33.3 (3)	19.4 (67)	.23* (70)	17.1 (70)

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

Storage — II

- Immunization Study illustrates importance of encoding
 - 70 parents interviewed as they immunization clinic
 - Reports compared to records
 - Never noticed vs. encoded incorrectly
 - Tended to recall same shots 10 weeks later
 - Number of shots related to accuracy

Comprehension

Comprehension Problems

- 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 Kinds of Ambiguity

(1) Lexical ambiguity:

The best way to prevent cancer is to catch it early

strongly agree somewhat agree somewhat disagree strongly disagree

- Unclear which sense of “catch” is intended
 - word label corresponds to more than one meaning

(2) Mapping ambiguity

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?

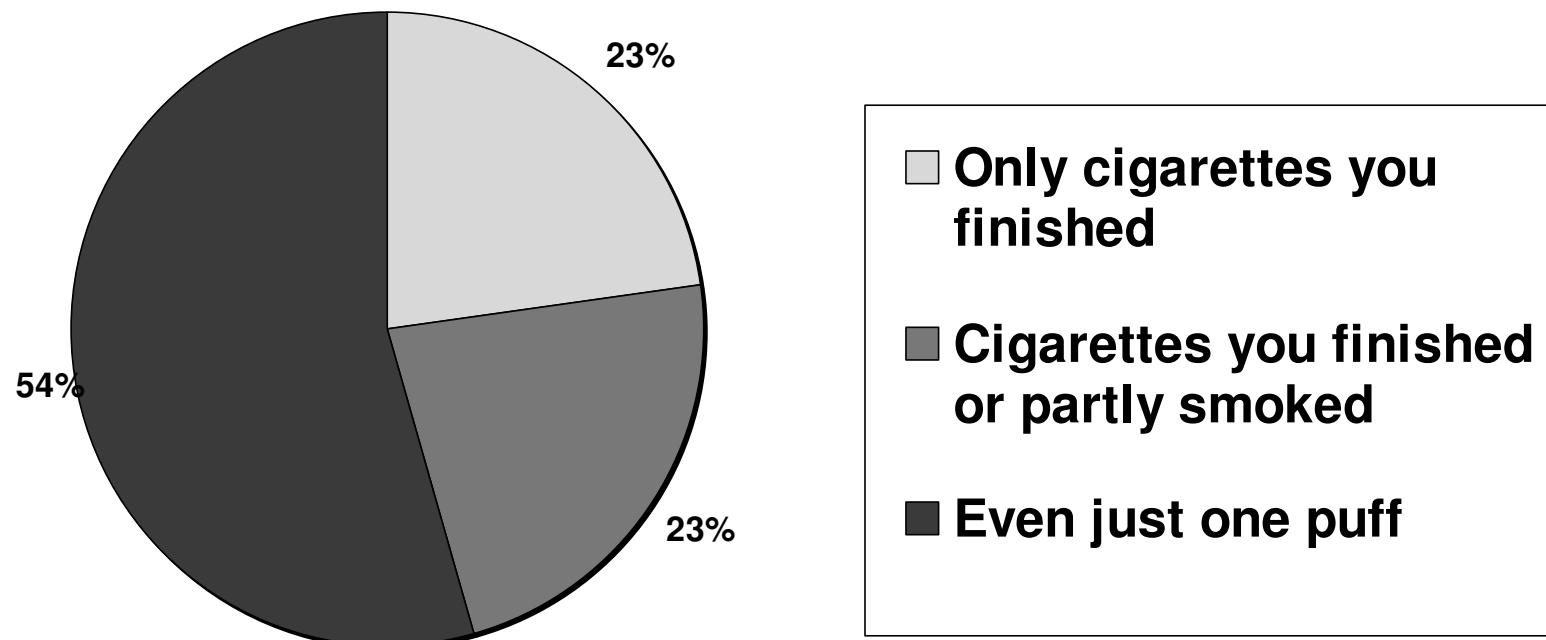
- Unclear how to apply “more than one job” to ones circumstances
 - which circumstances are included and which are not?

Two Kinds of Solutions

- In first case
 - pretesting, e.g. cognitive interviewing, should uncover problem
 - can repair with less ambiguous question wording
 - e.g. replace “catch” with “detect”
- In second case
 - pretesting cannot anticipate all borderline circumstances, especially if there are a lot
 - question cannot be worded to address all of them
 - can repair with clarification during interview
 - e.g. “in this survey we count that as one job”

Comprehension Problems: Conceptual Variability

- 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%



Comprehension Problems: Conceptual Variability

Do you think children suffer any ill effects from watching programmes with violence in them, other than ordinary Westerns?

- Belson (1981) determined that Rs interpreted *children*, *ill effects* and *violence* in numerous ways
 - e.g. “children”: < 8 years, < 19 - 20 years
 - children as students
 - only 8% interpreted question as intended
- Additional words can clarify intended meaning but this may lead to unwieldy questions
 - tradeoff between clarity and complexity

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.

Comprehension Problems: Presupposition

- Understanding a question involves recognizing both presupposition and focus of question
- What if presupposition is not accepted by R?

Family life often suffers because men concentrate too much on their work. Do you strongly agree, somewhat agree, etc.?
- “Don’t know” option does not solve the problem because implies acceptance of presupposition?

Comprehension Problems: Vague quantifiers

- Non-numerical terms for quantity have different numerical interpretations
 - Belson (1981) found “few” (*in over the last few years*) meant:
 - “no more than two years” (7/59)
 - “seven or more years” (19/59)
 - “ten or more” (11/59)
- Particularly problematic in response options

How often do you feel really excited? Very often, pretty often, not too often or never?

 - If R says more than *never*, Schaeffer & Bradburn asked for number
 - For educated and younger Rs, “pretty often” and “very often” were associated with larger numbers

False Inferences: Gricean Implicature

- Grice: Philosopher looking at why everyday use of words, sentences differs from logic: The cooperative principle: Conversations governed by cooperative principle; it's in everyone's interest to make conversation work smoothly
- Grice's four maxims:
 - Make your contribution as informative as needed
 - Speak the truth
 - Be relevant
 - Be clear

Gricean Implicature

- Person says something; listener fills in gaps (“implicatures”) that make sentence appear cooperative

“I’m out of gas”

“There’s a gas station around the corner”

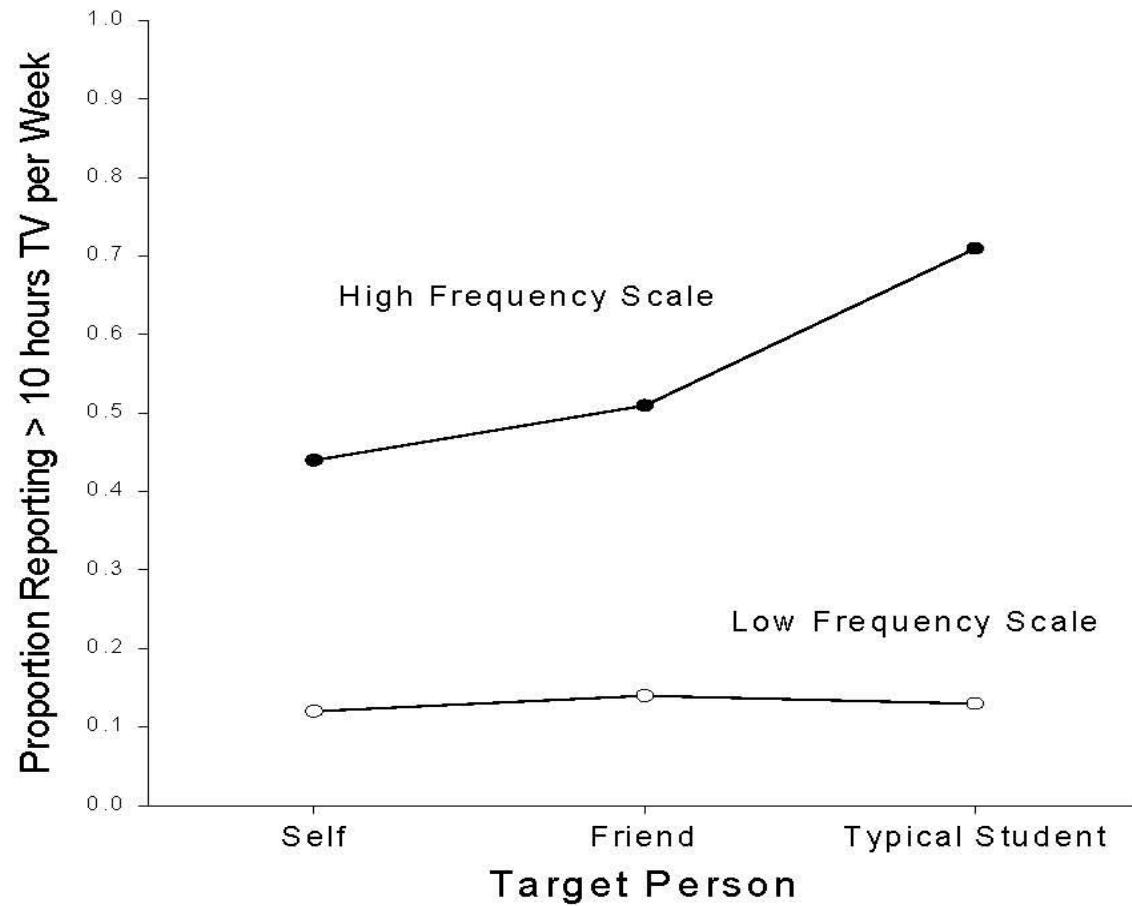
For this response to be cooperative, station should be open

- Utterance meaning is highly context dependent
 - “It’s raining out” could mean “Oh no, the picnic will be cancelled” or “Grab the book you left outside” or “Bring the umbrella”
- Problems arise if respondent assumes that a particular feature is intended to convey implicature; respondent incorrectly assumes that feature follows maxim of relation (“Be relevant!”)

Response Alternatives and Television Watching (from Schwarz & Bienias, 1990)

Low Frequency Scale	High Frequency Scale
() up to 2 ½ hours	() up to 10 hours
() 2 ½ to 5 hours	() 10 to 15 hours
() 5 to 7 ½ hours	() 15 to 20 hours
() 7 ½ to 10 hours	() 20 to 25 hours
() more than 10 hours	() more than 25 hours

Schwarz & Bienias—Results



Forgetting and Other Memory Issues

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
- Conrad, Brown & Dashen (2003) found recall better when asked about *taxonomic categories* than *properties* of events
 - Rs study list of ordinary nouns (e.g. *rose, Detroit, blood, chair, violet...*)
 - Assume Rs will spontaneously encode category membership but not properties
 - asked either how many instances presented from taxonomic categories (e.g. flowers) or with property (e.g. red)
 - r (actual and estimated) = .75 for taxonomic, .36 for property

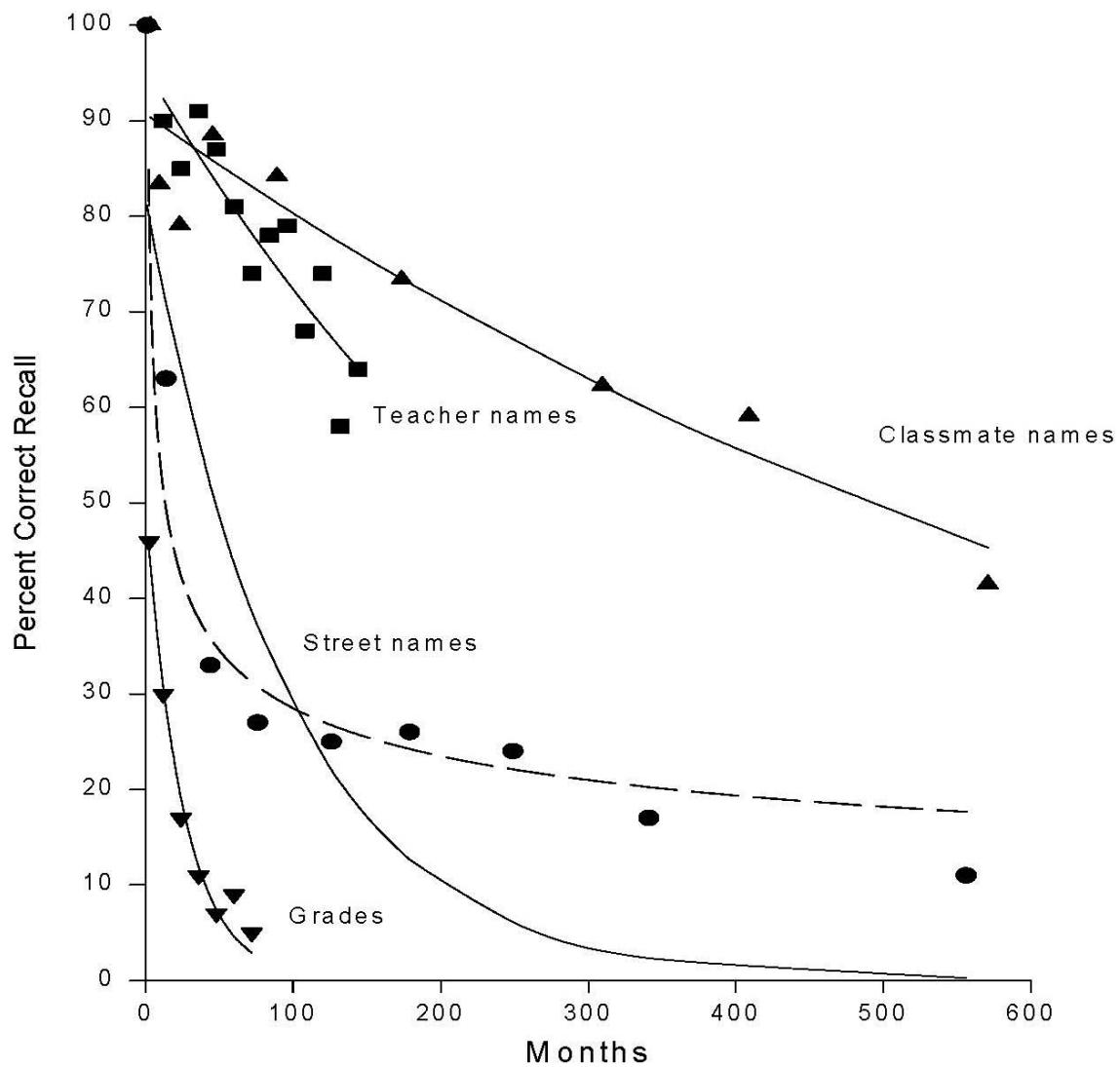
Memory Problems: Distortions Over Time

- 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



Retention Curves

- 1) Exponential decay--retain a fixed proportion in each unit of time:

$$\% \text{ recalled} = a e^{-bt}$$

- 2) Hyperbolic--event lost among similar events:

$$\% \text{ recalled} = 1 / (a + bt) \quad -b \text{ rate of accumulation of similar events}$$

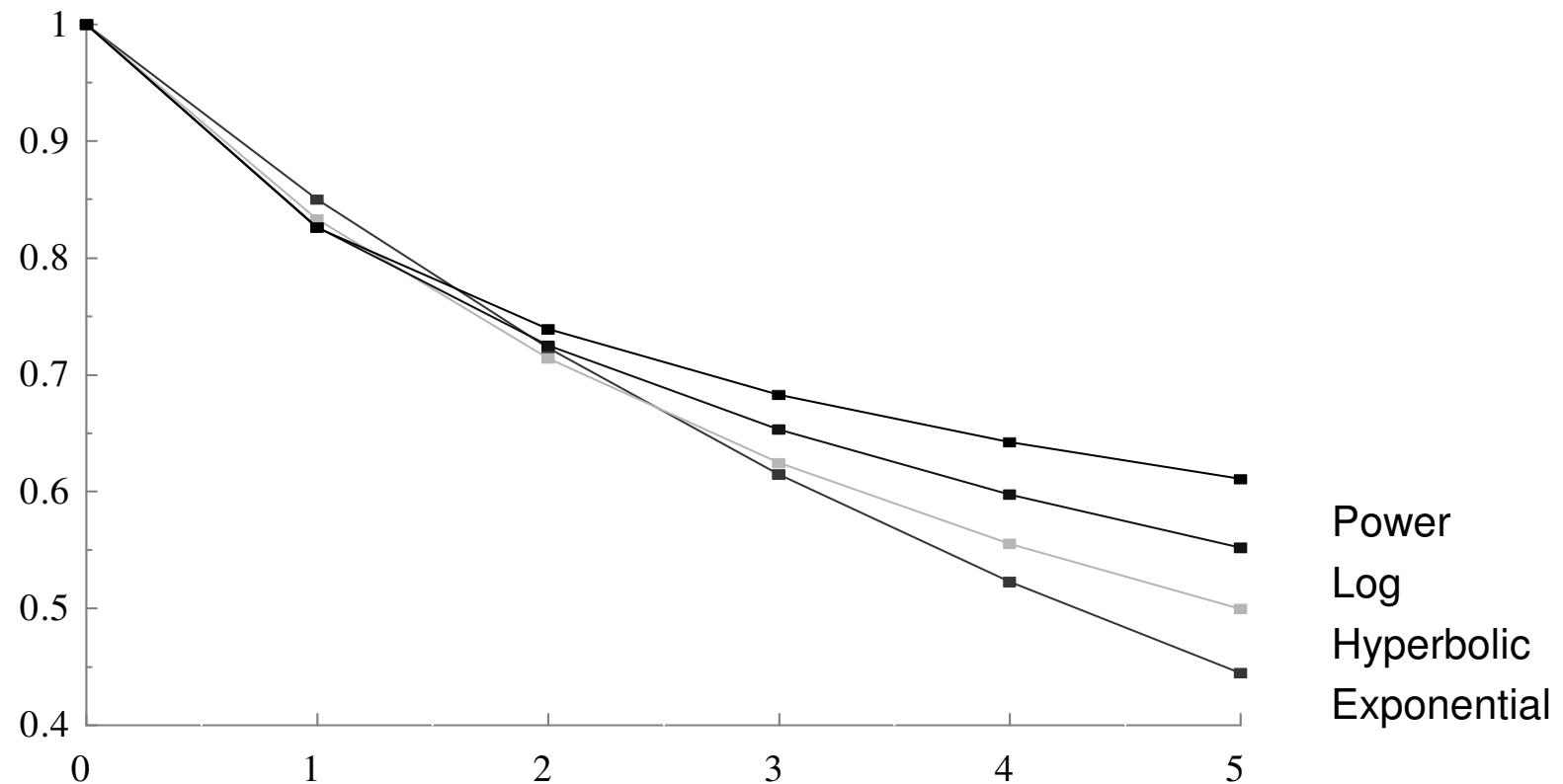
- 3) Logarithmic--equal ratios of time produce equal amounts of loss:

$$\% \text{ recalled} = a + b \ln (t+1)$$

- 4) Power -- equal ratios of elapsed time produce equal ratios of recall:

$$\% \text{ recalled} = a / (t+1)^b$$

Theoretical Curves



Impact of Cues

- What do we remember? What cues work best? Wagenaar's study

Cue	To be Recalled Item			
	What	Who	Where	When
What	—	62%	58%	20%
Who	26%	—	28%	7%
Where	29%	29%	—	13%
When	2%	2%	3%	—

Event and Question Characteristics

- Event Characteristics
 - Time of occurrence
 - Proximity to temporal boundaries
 - Distinctiveness/uniqueness
- Question Characteristics
 - Fit and specificity of cue
 - Time on task
 - Decomposition

Reconstruction Errors

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

Impact of Telescoping

Study	Bounding Procedure	Ratio of Events Reported: Unbounded over Bounded	
Neter & Waksberg (1964)	Prior interview	Expenditures Jobs	1.40 1.55
Loftus & Marburger (1983)	Experiment 1	Landmark event	Any victimizations 6.15
		Landmark event	Victim of theft 1.51
			Victim of assault 1.52
	Experiment 3		Reported crime 1.22
		Personal landmark	Any victimizations 5.50
	Experiment 4	New Year's Day	Any victimizations 2.00
Experiment 5	New Year's Day Specific Date	Any victimizations	2.52
		Any victimizations	1.32

Landmark Events for Bounding

- Gaskell et al. paper examine two “landmark” events
 - Thatcher’s resignation as PM
 - Hillsborough football disaster
- Neither event placed very accurately in time (15% within one month on Thatcher, 10% on Hillsborough)
- Direction of telescoping errors (predominantly forward errors with Thatcher, backward errors with Hillsborough)
- Are landmarks *not* useful?

Judgment and Estimation

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

Summary: Main Strategies for Frequency Questions

Strategies for Answering Frequency Questions	
Type of Information Used	Strategy
Information about Specific Events	<i>Recall-and-count</i> (episodic enumeration): Recall each event and count them to get the total number <i>Recall-and-count by domain</i> : Recall and count events separately by domain <i>Recall-and-extrapolate</i> (rate estimation): Recall a few events to estimate a rate and then project that rate over the reference period
Exact Tally	<i>Tally</i> : Recall current tally of events
Generic Information	<i>Retrieved rate</i> : Retrieve existing information about rate <i>Recommended rate</i> : Retrieve information about the recommended rate and then adjust upward or downward
General Impression	<i>Guess</i> (rough approximation; direct estimate) <i>Context-influenced estimate</i> : Use the value given by the middle response category as an anchor and adjust based on impression

Relative Use of the Different Strategies

Strategy	Study			
	Blair & Burton (1987)	Willis et al. (1999)		Brown & Sinclair(1997)
		Study 1	Study 2	Conrad et al. (1998)
Recall of Specific Information				
Recall-and-count	27%	27%	24%	42%
Other	3%	48%	68%	—
Estimate Based on Generic Information				
Rate	53%	—	—	2% 15%
Other	9%	11%	3%	— 9%
Tally	—	—	—	28% —
Impression	5%	11%	5%	14% 18%
Other/Uncodable	3%	3%	—	14% 18%

Judgment Problems: Opinion Questions

Now turning to business conditions in the country as a whole -- do you think that during the next 12 months we'll have good times financially, or bad times, or what?
(SCA)

- Assuming R has not already formed an opinion, similar set of strategies
 - vague impression
 - construct evaluation (either from general values or specific beliefs)
- Harder (or maybe impossible) to assess the truth of attitude than behavioral reports

What are Answers 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

Reporting: Formatting and Editing

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 unwillingness to be precise because truthful response is embarrassing

Range and Frequency

- Parducci's (1965) *range-frequency* principle suggests that survey respondents
 - match the categories to successive segments of their judgment (opinion, vague quantifier, etc.) -- the range
 - E.g. identify extreme examples and use as scale anchors
 - use the categories equally often -- the frequency principle
 - If stimuli are clustered on objective dimension, Rs tend to spread out so that can assign to different categories
 - uninformative to assign all stimuli to one category
- Parducci: Rating is average of values produced by two principles

Rating Scales (2)

- Other processes may modify initial representation
 - Schwarz & Wyer ('85) found that if previous rank ordering is from most to least, subsequent ratings shift toward “most” end of scale -- even if substance of tasks is unrelated
 - anchors scale as in Tverksy and Kahneman’s ('74) heuristic
 - Rs focus on end of rating scale presented first in ranking scale and make finer discrimintions
 - *positivity (or leniency) bias* leads to more frequent choice of positive options

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

not at all |-----|-----|-----|-----|-----|-----|-----|-----| extremely
successful
successful

0

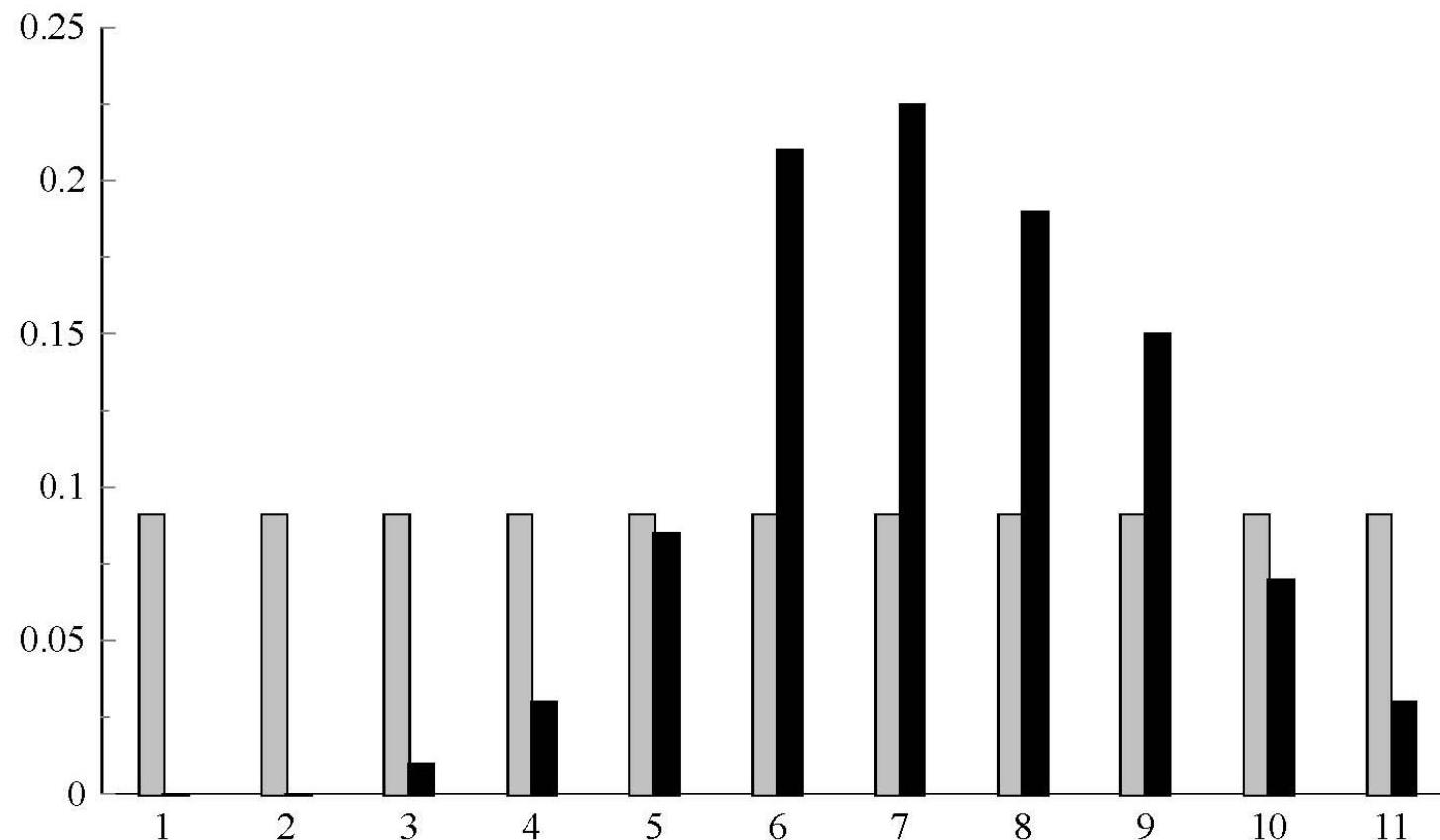
10

-5

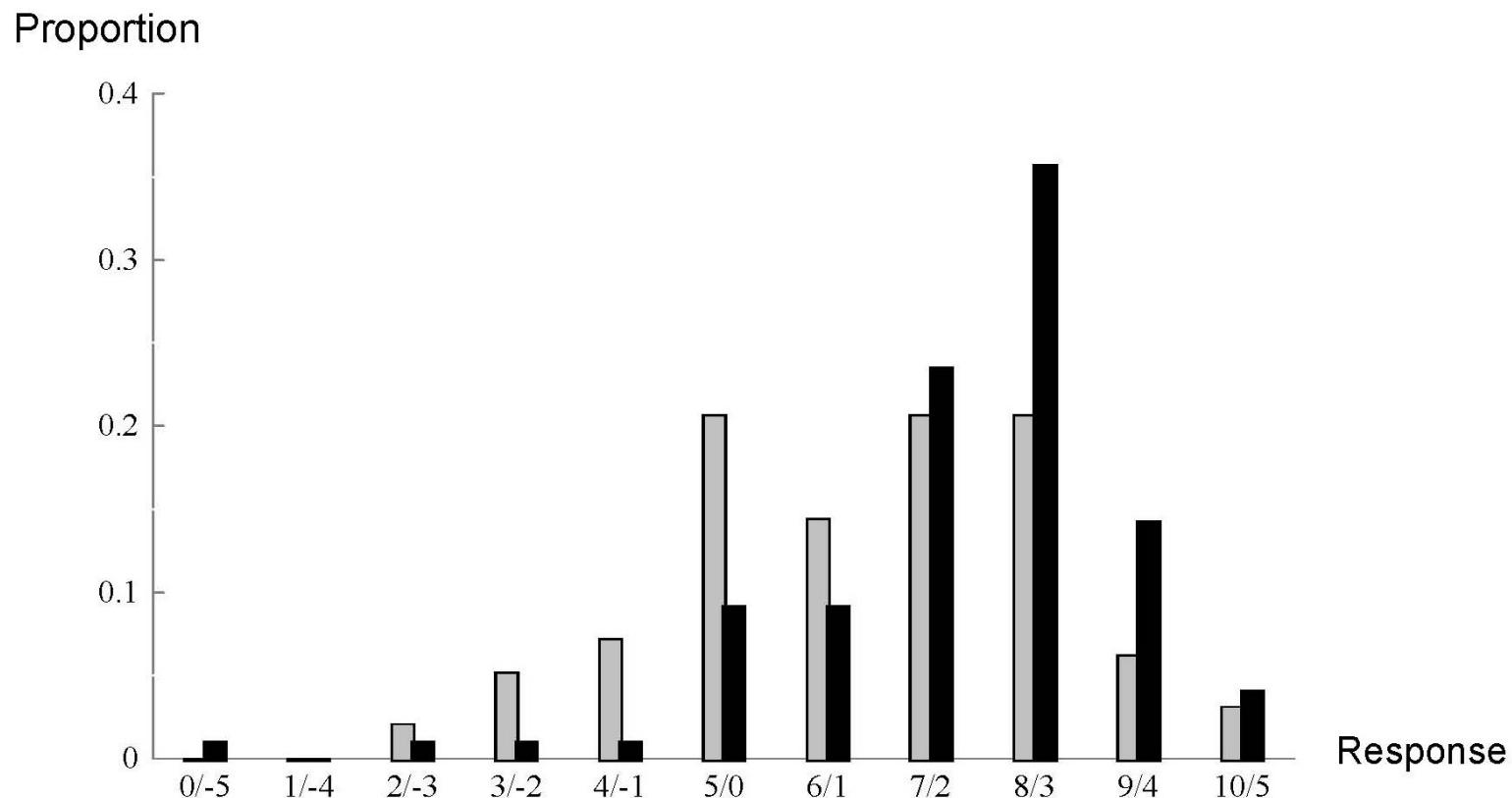
+5

The Frequency Principle and Positivity

Proportion



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

Editing

- Deliberate alteration of response
- Three main reasons
 - Consistency
 - Deference
 - Social Desirability

What's Sensitive?

- Three separate meanings
 - Intrusiveness: Question is inherently offensive
 - Social desirability
 - Some questions call for the respondent to provide information on topics that have highly desirable answers ... If the respondent has a socially undesirable attitude or if he has engaged in socially undesirable behavior, he may ... desire to appear to the interviewer to be in the socially desirable category. It is frequently assumed that most respondents resolve this conflict in favor of biasing their answer in the direction of social desirability.
 - Depends on topic, presence of interviewer, facts of the case
 - Disclosure to third parties

What's Sensitive? (cont'd)

- Respondent may not want to disclose information to researcher, to interviewer, to outside parties
- Which topics?

Topic	% Very Uneasy
Masturbation	56.4
Use of Marijuana	42.0
Sexual Intercourse	41.5
Intoxication	29.0
Petting and kissing	19.7
Income	12.5
Drinking	10.3

Consequences of Sensitivity

- Nonresponse: People fail to participate at all
- Missing Data: They skip offensive, embarrassing questions
- Reporting Errors: They overreport and underreport

Privacy Concerns and Nonresponse

- Confidentiality promises in federal surveys
 - As a matter of policy, the National Center for Education Statistics is required to protect the privacy of individuals who participate in voluntary surveys. We want to let you know that:
 - Section 406 of the General Education Provisions Act (20-USC 1221e-1) allows us to ask you the questions in the questionnaire.
 - You may skip any questions you do not wish to answer.
 - We are asking you these questions in order to gather information about what happens to students as they leave high school and make decisions about postsecondary education and work.
 - Your responses will be merged with those of others, and the answers you give will never be identified as yours.
- Meta-analysis: small effects (can backfire when topic not sensitive)

Participation in the Census

- Studies by Singer and her colleagues of “nonresponse” in the 1990 census
- Measured privacy (*The census is an invasion of privacy*) and confidentiality (*People’s answers to the census cannot be used against them*) attitudes
- Some evidence for impact of both privacy and confidentiality attitudes

Missing Data

Topic	% Answering
43 Items on a range of topics	97.3
Five demographic questions	98.7
Nine questions on illicit drug use	98.6
Eight questions on sexual behavior	97.9
Six questions on sexually transmitted diseases	97.4
Data from Tourangeau et al., 1997	

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
- Two key papers by Bradburn and colleagues
 - Blair, Sudman, Bradburn, and Stocking (1977)
 - Locander, Sudman, & Bradburn (1976)

Impact of Self-Administration

- Abortion (Lessler & O'Reilly, 1997; London & Williams, 1990; Mott, 1985)
- Illicit Drug Use (Many papers)
- Teen Smoking (Brittingham, Tourangeau, & Kay, 1998)
- Church Attendance (Presser & Stinson, 1999)

Impact of Open-Ended Questions

- Schwarz, Hippler, Deutsch, and Strack (1985)

Low Options		High Options	
Response Options	% Reporting	Response Options	% Reporting
Up to ½ hr.	7.4	Up to 2½ hrs.	62.5
½ to 1 hour	17.7	2½ to 3 hour	23.4
1 to 1½ hours	26.5	3 to 3½ hours	7.8
1½ to 2 hours	14.7	3½ to 4 hours	4.7
2 to 2½ hours	17.7	4 to 4½ hours	1.6
> 2½ hours	16.2	> 4½ hours	0.0

Impact of Open-Ended Questions (Cont'd)

- Tourangeau and Smith (1996)

Format	Mean Partners (5 Years)
Closed Low (0, 1, 2, 3, 4, 5+)	2.62
Open	3.12
Closed High (0, 1-4, 5-9, 10-49, 50-99, 100+)	5.33

Recent Meta-Analysis: Impact of Computer Assistance

- Tourangeau and Yan (2007) find no clear pattern
 - Mean effect size: 0.08
 - Standard error: 0.07

Conclusions—Editing

- Misreporting gets worse as items get more sensitive
- Misreporting worse among those with something to hide
- Four things help:
 - Self-administration
 - Randomized response technique
 - 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

Conclusions

Cognitive Approach (CASM) as General Theory

- Part of the more general movement to put survey research on a scientific footing
- Survey sampling already has that kind of foundation—laid out in Cochran, Kish, Deming, and especially Hansen, Hurwitz, and Madow (1953): statistical model of survey error
- General consensus among survey research that that model played out: O’Muircheartaigh: “Every time a new effect is found, we add a new variance component”
- Need for analysis of *causes* of error (nonresponse), built on social science theories (e.g., theories of persuasion)

Tourangeau (1984) as Guiding Paradigm

- Problem 1: The model needs to be specified in more detail; different processes relevant to different types of question
 - With behavioral frequency questions, starting to see better theoretical development, more specific predictions
 - With most types of questions, though, processes only vaguely specified
- Problem 2: The theory is too narrow; application of a certain type of cognitive analysis, overemphasizes comprehension and retrieval issues
 - Ignores cross-cultural/subcultural variation
 - Ignores interpersonal aspects of interview; Suchman and Jordan, conversational analysis as alternative paradigm
 - Says little about motivation (Krosnick's satisficing model a little stronger in this regard)
 - Fails to take into account alternative cognitive paradigms (PDP)

Other Problems with CASM Paradigm

- Problem 3: Too much explanation, too little prediction
- Problem 4: Every time a new effect is found, we add a new component to the response process
- Good paradigms need to be big enough to accommodate new findings but not so flexible as to completely avoid falsifiability
- Two other example of importation to social science models:
 - Use of models of social influence to understand nonresponse
 - Conversational analysis as tool for understanding respondent-interviewer interaction

CASM as Two-Way Street

- Scientific contributions from CASM researchers to areas of psychology
- Three areas:
 - Frequency estimation
 - Context effects
 - Memory for dates

CASM as Clinical Practice

- CASM conceived of as an applied enterprise; ideas, methods to be taken from elsewhere (psychology) and applied to survey problems (measurement error)
- Two general classes of innovation introduced under CASM banner
 - Response aids
 - Pretesting methods

Conclusions

- CASM has altered how survey researchers think about survey measurement error
- Also has highlighted measurement error as source of error
- Has yielded modest practical improvements, such as life events calendars and use of landmark events
- Most dramatic change—use of cognitive interviewing and other pretesting methods—in many ways least well supported
- Wide but superficial acceptance of new approach