Using Decomposition as a Questioning Strategy to Improve the Accuracy of Reporting about Events and Behaviors

Jennifer Dykema
Nora Cate Schaeffer
University of Wisconsin Survey Center
University of Wisconsin-Madison
January 12, 2006

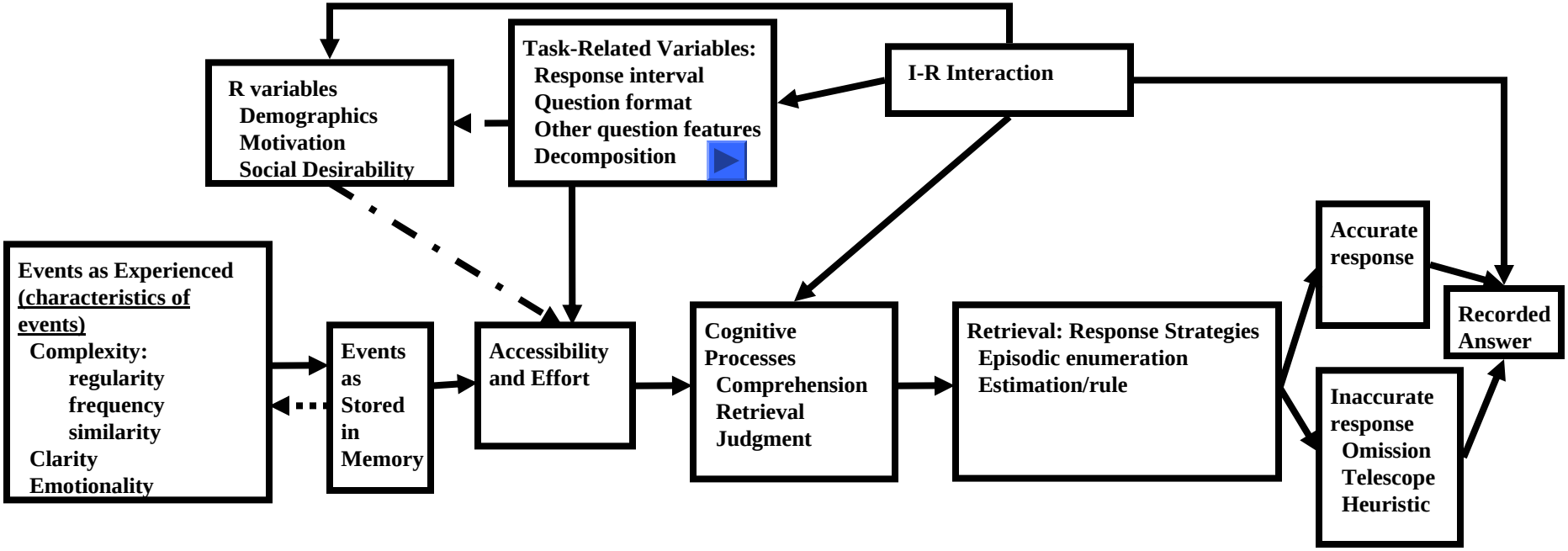
University of Wisconsin Survey Center

www.uwsc.wisc.edu

Assessing Behavioral Frequencies: Central Research Questions

- Exchanging child support payments is used as a case study to address the following questions:
- To what extent do the structural features or characteristics of a behavior predict inaccuracy in reports about the behavior?
- Can the analysis of errors generated by these characteristics be used in questionnaire design to improve response accuracy?
- Can questions about behavioral frequencies be designed to reduce the effects that behavior's characteristics have on reporting inaccuracy.

Framework: Characteristics and Processes Involved in Reporting a Behavioral Frequency



Decomposition

- **Definition:** division of single, global frequency question into two or more mutually exclusive, less cognitively taxing sub-questions; cuing along a relevant dimension; recall/estimation of frequency for the sub-questions; and combination of the sub-questions
- **Arguments in favor**: improves clarity; improves retrieval
- **Arguments against**: have to use right cue; category-split effects; could impair recall if best-remembered using a rate
- **Empirical evidence**: contradictory; may encourage overreporting

Data and Sample

- Validation Data: Court Record Database (CRD)
  - Based on abstracts from court records of all child support payments that were recorded by the courts
  - Contains demographic information on the parents at the time of the petition.
  - Provides valid measures of support-related variables.
- Survey Data: PS1, PS2, PS3
  - All surveys use the CRD as their sampling frame.
  - Telephone interviews.
  - Different questioning techniques to measure support.

PS1 Questioning Strategy

J.69. How often were the regular child support payments you were supposed to receive missed completely in 1986? In 1986 were all payments missed completely, some payments missed completely, or were all payments made?

J. 70. Were payments missed completely more than 6 times, 3 to 6 times, once or twice?

J. 71. What was the dollar amount of the child support payments that you received in 1986?

- Based on the NSFH.
- If all payments missed, $0 is imputed; if all payments made, amount paid is imputed from amount owed so uses owed to anchor reports.
- Cues frequency dimension but not similarity.

PS2 Questioning Strategy

S45. In (NAME EACH MONTH NOT ACCOUNTED FOR) did you receive any child support payments?

S46. In which months did you receive child support payments?

S47. Please tell me the amount of each payment.

- Based on analysis of response errors for the PS1.
- Structured decomposition using temporal cue of months.
- Cues frequency, regularity, and similarity dimensions by asking for amount of each monthly payment.
- Good strategy for those who exchange support on a monthly basis.

PS3 Questioning Strategy

- Based on CPS-CSS but informed by analyses of response errors for the PS1 and PS2, and from respondents' reports about recall for the event in cognitive interviews.
- Attempted to improve clarity by asking about support received through welfare before asking about regular child support.
- Experimented with filter question about exchanging any support by asking about receiving "even one" payment.
- Used a structured decomposition and a semi-structured to measure support exchanges among remaining respondents.
- Decompositions explicitly cue frequency and similarity dimensions and implicitly cue regularity.

PS3 Questioning Strategy: Structured Decomposition

328. During 1996, did you actually receive all of the (METRIC: weekly/every other week's/twice monthly/monthly/quarterly/yearly) child support payments you were supposed to receive for the children?

328a. How many (METRIC) child support payments did you actually receive in 1996?

328b. You said earlier that each payment was supposed to be [AMOUNT] dollars. Were all of the (METRIC) child support payments that you actually received in 1996 for the full [AMOUNT] dollars?

328c. Were (all of the (METRIC))/both of the/all of the [NUMBER]) child support payments that you actually received in 1996 for the same amount, or were (they/some of the child support payments) for different amounts?

328d. What was the amount of (this/each of these) child support payment(s)?

329a. So you received [AMOUNT] dollars (METRIC) in child support in 1996. That makes a total of [TOTAL] dollars in 1996. Is this correct?

PS3 Questioning Strategy: Semi-Structured Decomposition

329b. I need your help figuring out how much child support you actually received altogether in 1996. This information may be difficult to remember so please take as much time as you need and try to be as accurate as possible.

((If you received all the payments you were supposed to, that would be (52 weekly payments./26 payments every other week./24 bimonthly payments./12 monthly payments./4 quarterly payments./1 annual payment.)) / Before you told me you actually received [TOTAL NUMBER OF PAYMENTS] payment(s).)

I'd like to record the number of payments you received and the amount of those payments, so we can calculate the total. For example, you might have received 1 payment for 50 dollars and 5 payments for 30 dollars each.

Could you tell me the number of payments you received in 1996 and their amounts?

__ # of payments __ amount __ subtotal __ total amount __ total payments

Summary of Descriptive Statistics for Measures of Response Errors in Reports of Child Support Payments, PS1, PS2, and PS3

	Proportion		Errors Logged		
Sample	Error = $0	Error From -$100 to $100	Absolute	Signed	Correlation SR & CR
All	PS3 Highest	PS3 Highest	PS3 Lowest	PS3 Lowest	PS2 Highest
0 Freq	PS3 Highest	PS3 Highest	PS3 Lowest	PS3 Lowest	---
Freq > 0	PS1 Highest	PS1 Highest	PS2 Lowest	PS2 Lowest	PS2 Highest
Freq = 10-12	PS1 Highest	PS1 Highest	PS1 Lowest	**PS3 Lowest**	PS1 Highest

Results in table reflect tendencies, not statistically significant differences.

- Different measures of error sometimes produce opposing results.
- No tendency for Rs exchanging support (CRD) to overreport (not shown here).
- PS3: better reporting for “All” and “0 freq” Rs; included “even one” in filter.
- PS3: very successful in reducing bias among Rs with high frequencies.

Predictions

- Rs should be less accurate when events are complex (moderately frequent, irregular, dissimilar), lack clarity (are indistinct from like events), and emotionally neutral.
- Indicators for complexity, clarity, and affective intensity should be better predictors of reporting errors than other factors found to be associated with reporting errors (e.g., memory decay, respondent characteristics, social desirability, respondent motivation).
- The effects of complexity and clarity in the PS3 should be reduced in comparison to the effects of these variables in the PS1 and PS2 as a result of the PS3 questioning strategy.

Error Analysis: Comparison Among Surveys

Independent Variables Complexity	Absolute Error Logged			Signed Error Logged		
	PS1	PS2	PS3	PS1	PS2	PS3
Frequency	↓	↓	↓	**U***	**U***	ns
Payment pattern [Reg]						
Semi-regular	↑	ns	↑	**U***	**U***	ns
Irregular	ns	↓	ns	**U***	ns	ns
Metric [Semi/Monthly]						
No metric	ns	↑	ns	ns	ns	ns
Metric changed	↑	↑	ns	ns	ns	ns
Weekly/Biwkly	ns	↑	ns	ns	ns	ns
# Changes in Amount	↑	↑	ns	ns	ns	ns
Multiple of constant	↓	↓	ns	ns	ns	ns

Error Analysis: Comparison Among Surveys

Independent Variables	Absolute Error Logged			Signed Error Logged		
Clarity	**PS1**	**PS2**	**PS3**	**PS1**	**PS2**	**PS3**
Some alimony [none]	↑	↑	↑	ns	ns	ns
Health Ins [rcvr/neithr]	ns	↑	ns	ns	ns	ns
Mortgage [none]	---	ns	ns	---	ns	ns
Other $ [none]	ns	↑	ns	**O***	**O***	ns
Soc Services [none]						
1-6 payments	ns	ns	ns	ns	ns	ns
7-12 payments	↑	ns	ns	**O***	**O***	ns
Retroactive order [not]	↑	↓	ns	**U***	**O***	ns
Direct payment [none]	↑	↑	ns	**O***	**O***	**O***

Error Analysis: Comparison Among Surveys

Independent Variables	Absolute Error Logged			Signed Error Logged		
	PS1	PS2	PS3	PS1	PS2	PS3
Affective Intensity						
[Satisfied/dissatisfied]						
High dissatisfied	ns	ns	⬇	**U***	ns	**U***
High satisfied	ns	ns	ns	ns	**U***	ns
Memory Decay						
Days between end of reference period and Iw	ns	ns	⬇	ns	**O***	ns

Error Analysis: Comparison Among Surveys

Independent Variables	Absolute Error Logged			Signed Error Logged		
	PS1	PS2	PS3	PS1	PS2	PS3
Social Desirability						
Female Iwer [male]	ns	↓	ns	ns	ns	ns
Delinquency [not]	↓	ns	ns	ns	ns	ns
R Motivation						
Ever refused [not]	↓	ns	ns	**O***	ns	**U***
Income miss [none]	ns	ns	ns	ns	ns	ns
Both parents [one]	ns	ns	↓	ns	**O***	ns
Iw top quartile [not]	ns	ns	ns	ns	ns	ns

Error Analysis: Comparison Among Surveys

Independent Variables	Absolute Error Logged			Signed Error Logged		
R Characteristics	**PS1**	**PS2**	**PS3**	**PS1**	**PS2**	**PS3**
Male [female]	ns	ns	ns	ns	ns	**O***
Age @ Iw	↑	↑	ns	ns	ns	**O***
Income (logged)	ns	ns	ns	**U***	ns	**U***
# Children	↑	↑	ns	ns	ns	ns
Education [LT HS]						
High school	ns	ns	ns	**O***	ns	ns
GT high school	ns	ns	ns	**O***	ns	**O***
Mom phy cust [not]	ns	ns	ns	ns	**O***	**O***

Conclusions

- Demonstrated across three surveys that the characteristics of events are associated with response errors in reporting about the event.
  - These characteristics (complexity, clarity, and affect) are better predictors of error than other factors (memory decay, respondent motivation, social desirability, and respondents' characteristics).

Conclusions

- Taking into account the structural characteristics of events – that is, cuing the structural dimensions for respondents using the decomposition -- provided a methodology for writing questions to measure child support:
  - that resulted in more accurate reporting about the event overall.
  - that resulted in reduced impact of the structural characteristics on response errors.

Criteria for Making Decisions

…among various measures; selecting one Q strategy over another; choosing a style of interviewing (personal vs. standardized).

Statistic of interest (see Fowler and Mangione 1990):

validity versus reliability; bias versus precision; accuracy versus variation

Effects on entire sample versus subpopulations;

Effects on response distributions:

increase proportion of responses known to be under-reported (e.g., Cannell et al. 1981)

increase proportion of socially undesirable behaviors (e.g., van der Zouwen et al. 1991)

decrease item nonresponse (e.g., lower proportion of DK responses in van der Zouwen et al. 1991)

Effects in bivariate and multivariate models

Replication across surveys

Generalization from survey to population

Effects on the collection of data

(e.g., longer iws and more specialized interviewers with conversational interviewing (Schaeffer and Royston 1999))

Ease of use

quality of the interaction (e.g., measures of Rs or Ints levels of frustration or burden) (Schaeffer and Royston 1999)