

Response to Interview Structure Issue Paper

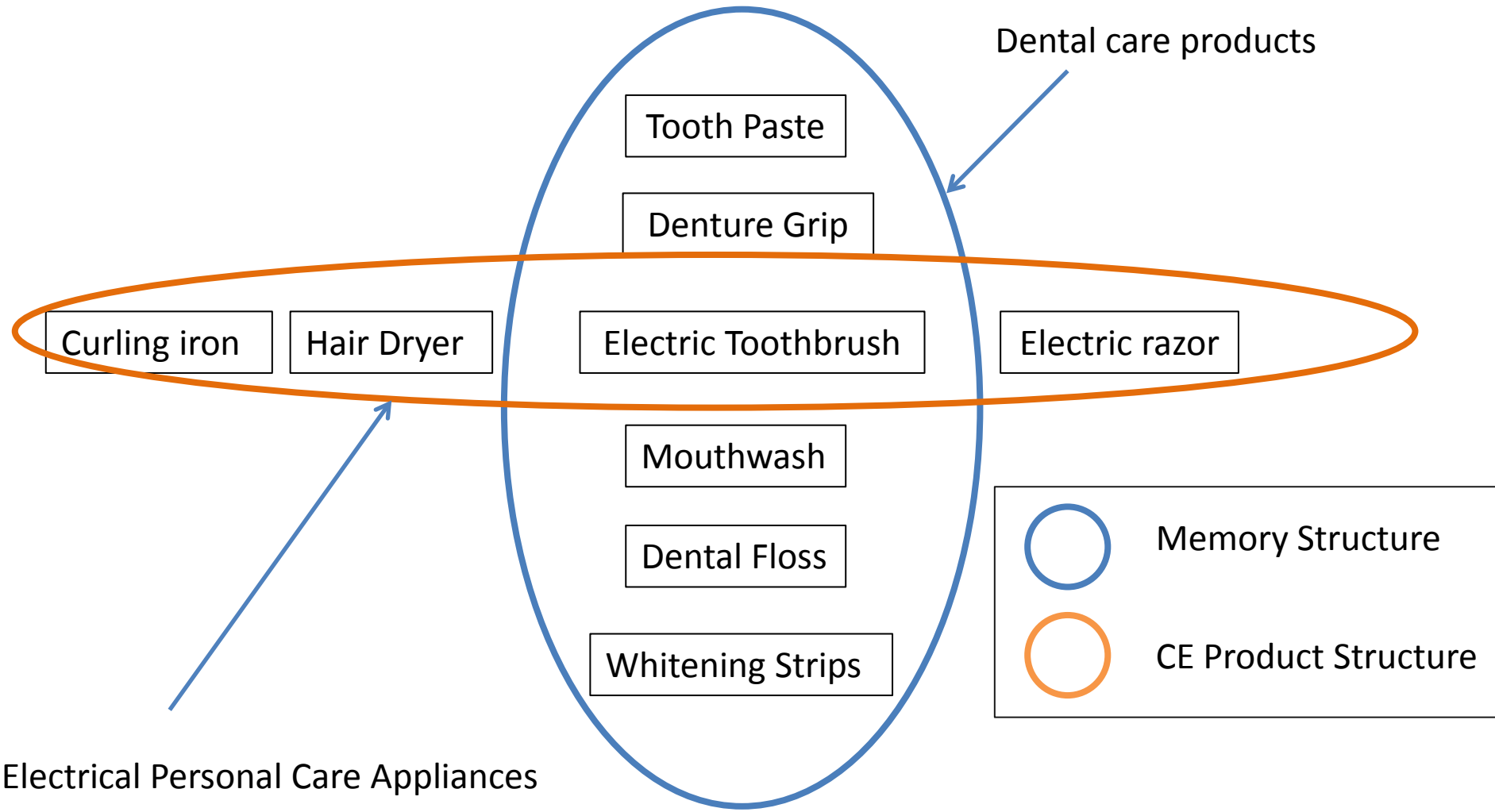
Frederick Conrad
University of Michigan

Why are expenditures are underreported?

- Main concern of “Interview Structure” issue paper is that expenditures are underreported
 - Stated more explicitly in Gemini Project Vision document, p.2
- It is proposed that
 - underreporting may be due to question order that cuts across *Rs'* memory for expenditures and so does not promote retrieval
 - a question order that better fits *Rs'* memory structures will promote more complete reporting

An Alternative Explanation

- Underreporting is not about the order of questions but the content of questions, in particular, the categories
- *Rs* may not think about their expenditure events, at the time they occur, as instances of the CE categories
- If they do not, then asking about CE categories in the interview won't bring relevant events to mind
 - irrespective of the order in which *Rs* is asked



Underreporting Instances of *Unnatural* Categories

- Conrad, Brown & Dashen (2004) tested this idea in lab experiment
 - Study phase: participants read 109 ordinary nouns, one at a time
 - Test phase: asked how many words just studied
 - belong to particular taxonomic category
 - contain a particular property
- Study phase analogous to purchase events
- Test phase analogous to CE interview
 - Taxonomic categories – natural categories – correspond to the “dental products” in the example
 - Properties – unnatural categories – correspond to “electrical personal care products”

Taxonomic Category Group

STUDY

dog
Chicago
guitar
maple
trout
violin
salmon
table
⋮

TEST

TREE
FISH
FURNITURE
TOOL
⋮

Implicit Property group

STUDY

corn
ammonia
chocolate
salt
garbage
ivory
peach
daffodil
⋮

TEST

SMELLY
YELLOW
FUZZY
ROUND
⋮

Explicit Property group

STUDY

YELLOW

corn

SMELLY

ammonia

BROWN

chocolate

WHITE

salt

SMELLY

garbage

WHITE

ivory

YELLOW

peach

YELLOW

daffodil

.

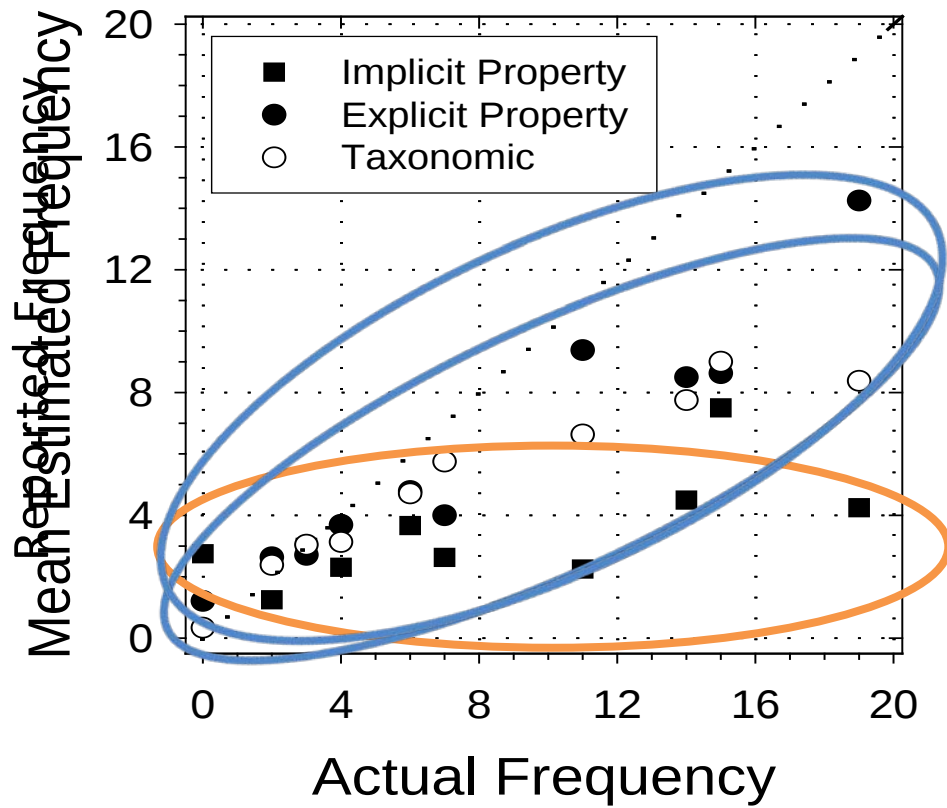
TEST

SMELLY

YELLOW

FUZZY

ROUND



Signed Error

Proportion "hits" to all
enumerated items

Implications of Study for CE Interview

- On the downside:
 - Asking *Rs* about expenditures from CE categories will not make contact with *Rs*' memory for many relevant expenditures, leading to underreporting
 - If there is misalignment during purchase episode (encoding), there may be little that can be done during the interview (recall) to help
- On the positive side
 - When *Rs* do encode expenditures as instances of CE categories, their recall will be good
 - Might be possible to train *Rs* across waves to think about expenditures as CE does

Making CE Categories Natural

- May be possible to intervene in first interview so that *Rs* learn to think of their purchases, when they occur, as instances of CE categories
- But probably too many categories to do this exhaustively
- Could focus on those categories most at odds with *Rs'* natural classification of expenditures
 - Would need to determine this with experiment like the one just described
 - May still be too many to train *Rs* on all and training might not be effective

Context Reinstatement

- An alternative to training *Rs* to think like CE analysts is help them recall all relevant purchases, irrespective of CE category, by helping the recall the purchase *context*
 - Since the last interview did you purchase anything online? On the phone? By mail? In drive-through outlets? In enclosed malls? In convenience stores? ...
 - If “yes,” what did you purchase?
 - Any reported purchases are coded into CE categories by interviewer or coder after the fact

Context Reinstatement (2)

- *R* might indicate she made many online purchases and then list “books,” “computer hardware,” “cell phone service,” “plane tickets” and “cosmetic surgery”
 - Note it is the *purchase* context not the *consumption* context that matters
 - *R* paid for air travel online but consumed the service in the air

Context Reinstatement (3)

- Used to improve eye witness reports as one of several memory improvement techniques known as *cognitive interviewing* (CI) (e.g., Bekerian & Dennett, 1993)
 - Not the same as pretesting method
 - Other memory improvement techniques in CI include varied physical perspective and varied recall order
 - context reinstatement probably locus improved recall in CI (Milne & Bull, 2002)
- Based on *encoding specificity* (Tulving & Thompson, 1973)

Context Reinstatement (4)

- In CI, participant generates context, e.g., of a crime, but idea for CE is to provide contexts to *R*
 - Should make it easier: more recognition than recall
- May be that contexts need to be presented at finer level of detail
 - e.g., “online purchases” may need to be decomposed into “online purchases that involved shipping a physical product,” “online purchases of a downloadable product,” and “online purchases of an offline service,” etc.

Similarity to Event History Calendars

- Proposed approach has some of the character of Event History Calendars (EHC) (e.g., Belli, Shaye & Stafford, 2001)
 - Recalled context stimulates subsequent recall
 - In EHC, recall from one life theme (e.g., employment) serves as cue for retrieval of events from another theme (e.g., residential moves)
 - Parallel retrieval (Belli, 1998)
 - In current proposal, interviewer provides context
- To the extent that purchases are narratives, extended over time, the context reinstatement and EHC approaches are similar
 - But this may not often be the case

Similarity to EHC (2)

- Question order:
 - EHCs inherently unstructured
 - context reinstatement approach noncommittal:
 - contexts must be presented in some order but no theoretical guidance (yet) on whether any order produces better recall than others
 - Certainly if recalling online book purchases brings to mind book purchases in brick and mortar outlet, sensible to record those purchases at that time

Flexible Data Entry

- Issue paper describes context effects as rationale for maintaining fixed question order
- Seems low risk for in this domain
 - Hard to see how asking about *home furnishings* before *clothing* for one *R* and reverse for another introduces substantial measurement error (underreporting)
 - especially if the different orders are the result of different self-generated reminders
- Makes sense to accommodate *Rs'* preference to report on one category (or context) by allowing interviewer to enter expenditures in whatever order *R* happens to report them

Flexible Data Entry (2)

- User interfaces to promote flexible entry
 - May require moving outside Blaise comfort zone
 - Representing questionnaire as clickable network would allow direct access to any question
 - Multimodal user interfaces could allow interviewer to
 - enter notes with stylus into onscreen notepad linked to categories
 - speak notes while entering data with keypad
- Johnston (2007) argues multimodal interfaces more natural than single mode for survey interviews because support everyday practice of combining speech, pointing, and gesture as needed

Conversational Interviewing

(e.g., Schober & Conrad, 1997; Conrad & Schober, 2000)

- Issue paper points to *conversational interviewing* as example of “order-free” interviewing
- I believe this somewhat mischaracterizes the approach that Schober and I have explored which is concerned with improving *R*s understanding of individual questions, not with variable question order
- The extra time required by the approach
 - due to the time taken to clarify question meaning and help *R*s establish the correspondence between question concepts and their circumstances
 - not variable question order

Research Program and Some Questions

1. *Unnatural Categories:*

- For which CE products are expenditures most underestimated under the current approach?
- In what contexts are they most likely to be purchased?

2. *Context Reinstatement*

- Does the proposed approach help Rs recover purchase events that do not come to mind when probed with CE categories?

3. *Flexible Data Entry*

- What are the temporal costs of following respondents' unstructured recall?
- Can interviewers do this effectively in real time?
- What user interface approaches, e.g., what combinations of input devices and modes, best support flexible data entry in semi-structured verbal tasks?

Thank You