

Asking Questions about Household Member Activities
to Improve Expenditure Reporting

Follow-Up to 2012 Proxy Reporting Lab Study

Erica Yu
Office of Survey Methods Research
September 20, 2013

Background

The Consumer Expenditure Quarterly interview survey (CEQ) relies on proxy reporting, a single respondent reporting expenditures on behalf of others, as a source of data. There are two known sources of error: knowledge and recall of others' expenditures (BLS Position Paper on Proxy Reporting, 2010; Mathiowetz, 1987). Respondents cannot report information they do not have and recall is often flawed even when the respondent is knowledgeable.

Previous research has found that the quality of proxy reporting varies as a function of several factors, including involvement in proxy purchases, communication style, and relationship type (Kojetin & Jerstad, 1997). However, these findings have been primarily focused on whether and how proxy information is encoded in memory. The current study aims to understand how to enable respondents to retrieve what proxy information is already in memory.

In 2012, Book and Edgar investigated ways to reduce the impact of the recall problem in proxy reporting. Focusing on errors of omission, Book and Edgar (2012) created a protocol designed to raise respondents' awareness of others' spending. Book and Edgar developed four questions that ask respondents to think about: the ways that others in their household spend money that they themselves do not, the hobbies and activities that others in their household like to do, the trips that were taken by others in their household, and any unusual purchases made by others in their household. When administering the CEQ, the interviewer used the information gleaned from the initial questions about others in the household to prompt the participant to think about any expenses that might have been forgotten. In response to the additional questions and probes, the 20 participants identified more than 20 items worth more than \$3,200 – which otherwise would not have been reported. These results suggest that additional survey questions designed to remind respondents' to include other household members may be effective at reducing underreporting of others' expenditures. However, Book and Edgar (2012) relied on highly-personalized questions, which require agility and familiarity with the CEQ categories on the part of the interviewer to maintain in memory and use effectively the provided responses throughout the survey. Non-scripted questions also require a high level of interviewer skill to be administered effectively. The current study develops standardized versions of the questions that may be more appropriate for field use.

While the evidence suggests that additional survey questions may reduce underreporting of others' expenditures, they also lengthen the survey duration. It is likely that the longer the interviewer spends with the respondent discussing household members' expenditures, the better able the interviewer will be to prompt the respondent for additional expenditures during the CEQ. Research suggests that respondents need several seconds to retrieve information (Reiser, Black, & Abelson, 1985) and that the amount of time allowed for participants to access events and the amount of time used to formulate responses can impact what type of response strategy is used and ultimately the type and quality of information that is retrieved in some recall

tasks (Blair & Burton, 1987; Burton & Blair, 1991). Longer response times are often associated with more accurate recalls while shorter response times are associated with inferior response strategies such as estimation. Implementing the proposed additional questions efficiently will require understanding the minimum time needed for participants to provide meaningful open-ended responses. The current study includes a timing manipulation to test the effect of interviewer pacing on responses to the proposed questions.

The primary aims of the current study are to refine the proxy reporting questions developed by Book and Edgar (2012), by standardizing script wording to be used in a production setting. A secondary goal of the current study is to understand the minimum amount of time needed to spend on these additional questions.

Method

Design of timing study

All participants received the same cognitive interviewing protocol except for one difference during the initial proxy questions: half of the participants were given a minimum of 10 seconds to complete their open-ended responses while the other half of the participants were given a minimum of 2 seconds. This design aimed to investigate whether additional time would enable participants to more thoroughly search memory for additional reports. The latter group aimed to approximate a style of interviewing wherein the interviewer continued on after the participant gave any response. A computer-assisted personal interviewing (CAPI) instrument was developed using Visual Basic to enable the interviewer to input summaries of the participants' responses, implement the 2 and 10 second manipulations, and record the time spent on each proxy question. When administering the 10 second minimum interval, the interviewer maintained normal interviewing style despite any silences by continuing to type or otherwise appear to be working until the minimum interval had passed.

Materials for proxy protocols

The initial proxy questions asked participants to consider how others in their household spend their money based on the four questions tested in Book and Edgar (2012) with several modifications. Question scripts and the full protocol are included in Appendix A. These questions ask participants to think of general, dispositional characteristics of others in their household (converging on a common reporting strategy that proxy respondents have been found to engage in; Schwarz & Wellens, 1997). A question about changes to day-to-day-routines was added to capture behavioral changes that are more likely to be recalled than expenditures but nonetheless result in expenditures (Reiser, Black, & Abelson, 1985). The addition of this question is further supported by evidence that household members commonly retrieve from memory information about proxy purchases by thinking of usual behaviors; a question asking about *changes* to routines may cue new information. Second, household member names were included in each question to further prompt participants to consider each household member. And third, the hobbies question was shifted to the end to reduce overlap with responses to the other questions due to its broadness. The open-ended proxy questions in the current study asked about:

- (1) types of things others spend their money on that the participant did not,
- (2) vacations or trips others had gone on without the participant,
- (3) changes to the day-to-day routines of others,
- (4) unusual or out of the ordinary purchases made by others, and
- (5) hobbies of others.

Follow-up probe questions were developed to be standardized and appropriate for field use. The original probe scripts are included in Appendix B. Probes were designed to remind the respondent of the information he or she had provided at the start of the interview about other household members' spending habits, trips, routines, unusual purchases, and hobbies. The interviewer followed closely the question scripting and did not use unscripted probes.

To preview the main finding of this report, however, the scripted probes did not elicit additional expenditures as expected. While the 2012 study found that 10 of 18 participants added at least one expenditure report after hearing a proxy probe, the current study found that 0 of the first 11 participants added expenditures. Given that the current study was intended to improve upon the 2012 study, the protocols were re-evaluated. The primary difference between the 2012 study and the current study was identified to be the scripting of the proxy probes used during the CEQ portion of the study. In the 2012 study, the interviewer asked conversational probes that were tailored to each participant's situation. The probes were not scripted in advance and varied across participants. One example was:
"You mentioned that your wife bought some school books, has she done that in the past three months?"

In contrast, the current study's original scripting dictated that the interviewer asks:
"You mentioned that your wife made a change to her day-to-day routine. Are there any other expenses that you want to add related to that?"

Despite the small size of the sample, the lack of additional proxy expenditure reports was deemed compelling enough to warrant modifying the probe scripts. Assuming that the current study's original probe scripts lacked sufficient cues to help participants recall proxy expenditures, the new probe scripts were designed to include more information from participants' initial open-ended responses while still maintaining standardization across participants. Using the same example, a modified probe script might read:
"You mentioned that your wife started school. Did she have any school-related expenses that we haven't talked about yet?"

The modified scripted probes were used for the remaining 14 participants; the protocol for each type of proxy report is included in Appendix B.

Following both the proxy protocol questions and the CEQ reporting sections, the participant was asked debriefing questions. The debriefing questions asked after the proxy protocols targeted the participants' comprehension of the proxy questions and knowledge of other household members' spending. The debriefing questions asked after the CEQ reporting sections targeted participants' judgments about reporting accuracy and value of the probes used. The scripted debriefing questions are included in Appendix C; additional unscripted debriefing questions were used as needed.

Materials for CEQ reporting sections

The CEQ was administered using the CAPI instrument used for training (v.14.43_05-24-2013). Training cases were selected from Wave 2 or Wave 3 and household rosters were modified to include at least two people.¹

¹ Despite using cases with multi-person households, scripts throughout the CAPI instrument prompted the interviewer to ask whether 'you' had any expenses, rather than the inclusive 'you or anyone in your

Participants

Twenty-five participants were recruited from the Office of Survey Methods Research (OSMR) database using two eligibility criteria: household size of at least two people (mean = 2.92 people, SD = 1.15) and the purchase of CEQ items (e.g., books, entertainment expenses, things for the home) during the reference period. Eligibility based on household size was further restricted to include people that were considered 'family' to target those households with shared expenses. The recruiter was also asked to select individuals across a range of education (Table 1), gender (number of male participants = 12), age levels (mean = 45.4 years, SD = 13.58; missing age data on 3 participants), income (Table 2), and varying roles in the household (e.g., parents with teenage kids, member of extended family; Table 3). The most common household types were participants living with children younger than 18 years (12 participants) and participants living with their partner only (6).

Table 1. Respondent education level (n=25)

<i>Education level</i>	<i>n</i>
Some high school	1
High school diploma	2
Some college	6
College degree	11
Advanced degree	5

Table 2. Respondent Income level (n=25)

<i>Income level</i>	<i>n</i>
Less than \$40,000	8
\$40,001-\$60,000	7
\$60,001-\$75,000	2
More than \$75,001	8

Table 3. Respondent household composition (n=25)

<i>Household type</i>	<i>n</i>
Adult and partner	6
Adult and child(ren)	6
Adult and extended family, mixed ages	5
Adult and roommates	4
Adult and partner and child(ren)	2
Adult and partner, visited by child for summer	1
Adult visited by child regularly	1

household'. Debriefings from early interviews confirmed that at least some participants thought primarily of their own personal expenses when hearing the 'you' prompt. Later interviews corrected this error.

Procedure

Eligible individuals were brought into the OSMR lab for interviews. Before beginning the interview, the interviewer gave a brief introduction to the purpose of the Consumer Expenditure Survey, the purpose of the study and the task and collected informed consent and permission to record audio. The interviewer then collected household composition, household income level, and respondent education level. Respondent age and gender are on file in the participant database.

The interviewer administered the proxy questions using a CAPI instrument. After this section of questions was completed, the interviewer asked debriefing questions regarding comprehension, sensitivity, and ease or difficulty of answering the questions. Where a response along a scale was requested, show cards were used.

The interview continued onto the CEQ portion of the study, in which the interviewer selected and administered relevant CEQ sections. The interviewer aimed to complete at least 3 sections, including at least one section indicated by the initial proxy-related questions and at least one section not indicated. Proxy-related sections were selected as categories that contained at least one of the responses provided during the proxy questions. As two examples: if the respondent said that a household member had a hobby of playing video games, Section 17 was administered; if the respondent reported that a household member had a change in their day-to-day routine because of a new job and so was spending more money on clothes, Section 9 was administered. Sections were chosen on-the-fly as those most appropriate for the participant's situation and thus not all participants were administered the same CEQ sections. At the start and end of each relevant CEQ section, the interviewer administered the relevant probes (Appendix B). After the CEQ portion was completed, the interviewer asked debriefing questions about the study overall.

Results

Effect of timing manipulation on the number of reports for proxy questions

Overall, participants reported a median of 6.0 items or activities across the 5 initial proxy questions (minimum=2.0; maximum=35.0). The overall mean number of items or activities reported is 7.3 (SD=6.9); however, this number is inflated due to one participant's 35 reports. The manipulation of minimum response interval was designed to enable participants more time to reflect on the question and generate their response. The data suggest the additional time may have led to additional reports: the median number of reports in the short interval group is 4.0; the median in the long group is 6.0. Given the small sample size, it is difficult to interpret such differences between the timing groups. Unless otherwise noted, the groups have been combined for all analyses.

Additionally, at least two participants initially answered "none" to a question and used the remainder of the minimum 10 second interval to continue thinking about the question and change their response. An interviewer using the standard pacing style would likely have moved on to the next question and missed collecting information.

Representation of other household members

Overall, for each of the five proxy questions asked, participants were able to provide information for the other members in their CU. The first question ("Are there some types of things that [name] spends money on that you don't?") was repeated for each individual in the CU. For this

question, 16 out of 48 consumer unit (CU) members across all participants were reported for as “none” or “nothing”, though this consists primarily of young children.

The four later questions (trips, routines, unusual purchases, hobbies) are asked at the CU-level though the question wording includes each CU member’s name and the information is recorded at the individual level. As shown in Table 4, a substantial proportion of participants responded simply “none” or “nothing” to questions except the question about hobbies. These high proportions may indicate that participants did not understand the question, did not find the question relevant to their household, or refused to answer.

Table 4. Number of respondents reporting no expenditures or behaviors (n=25 for each question)

<i>Proxy question</i>	<i>n</i>
Trips without participant	12
Changes in routines	11
Unusual purchases	16
Hobbies	2

Expenditure categories identified during proxy questions

Responses from the proxy questions may indicate the typical types of expenditures that respondents are able to recall and report for. The first proxy question (“Are there some types of things that [name] spends money on that you don’t?”) elicited the expenditure categories shown in Table 5 (a participant contributes at most one time to each category, regardless of the number of CU members indicated in response to the questions). The most common category was clothing, with 13 out of 25 participants naming clothing as an expense of at least one CU member.

Table 5. Categories generated in response to open-ended question about how others spend money

<i>Category named</i>	<i>n</i>
Clothing and accessories	13
Movies, books, and entertainment	7
Food	7
Vehicle expenses	5
Rent, mortgage, insurance, and utilities	5
Cosmetics and personal care	5
Toys	3
Electronics and video games	3
Transportation	2
Exercise and recreation equipment	2
Travel expenses	1
Home goods	1
Medical expenses	1

Thirteen participants reported that at least one CU-member went on a trip without them. Destinations included California (2 participants), Louisiana, West Virginia, and Arkansas; the beach (3); a cruise (2); going back and forth to college; Europe, Dominican Republic, and Switzerland.

Eleven participants reported that at least one CU-member experienced a change in their day-to-day routine. Responses included health conditions (3), family vacations (3, not included as response to the trips question), getting out of school for summer (2), summer camp, purchasing a vehicle, joining the gym, and canceling a phone plan.

Nine participants reported that at least one CU-member had made an unusual purchase. These purchases included trip souvenirs (2), jewelry, plane tickets, handbags, a car battery, toys, a DVD, and a stun gun.

Twenty-three participants reported a hobby for at least one CU-member, as shown in Table 6. The most common category reported was participating in sports or recreation. Additional categories named by only one participant include cars, watching sports events, shopping, and going to museums.

Table 6. Categories generated in response to open-ended question about others' hobbies

<i>Category named</i>	<i>n</i>
Sports and recreation	16
Electronics/video and computer games/surfing the internet	9
Music/movies/television	4
Reading/writing/studying	4
Family and friends	2

The five proxy questions collected a range of information. Importantly, respondents' answers spanned across many CEQ sections, with breadth and relevance that could not have been achieved with more specific questions. Not all responses were directly related to an expenditure category, but all reports could be used to prompt participants for additional expenses. In this way, probes are relevant to households even while the questions retain standardization across the sample. And, while the expenditure categories identified after probing may indicate which categories respondents are likely to be unable to proxy report for, it is also highly likely that there remain yet more expenditures not recalled at all.

Proxy protocol sensitivity

During debriefing, participants rated how sensitive the proxy questions felt on a scale from 1 to 5, where 1 represented not at all sensitive and 5 represented very sensitive. The most frequent response was that the questions were 'not at all sensitive', with the majority of responses falling along the lower range of the sensitivity scale, as shown in Table 7 (median = 2). However, when asked in a follow-up question to speculate about how others might feel, participants revealed that they believed the nature of the questions was potentially sensitive (median=3). This difference between the self and other ratings suggests that the answers, rather than the question, are the sensitive element. It is likely that participants offered no sensitive answers but

could imagine others revealing personal information. Indeed, several participants remarked that others' expenses were "not her business" or that providing information was "giving out information that was not hers to give." One participant reported that the questions made him feel like he ought to know the answers to the questions and so he had a sensitive feeling when he did not know an answer.

Table 7. Sensitivity ratings for proxy questions (n=25)

<i>Rating</i>		<i>Self</i>	<i>Others</i>
		<i>n</i>	<i>n</i>
1	Not at all sensitive	8	5
2		6	2
3		6	8
4		4	6
5	Very sensitive	1	4

Proxy protocol awkwardness

Given the inclusion of a potential silence in the interviewing protocol through the minimum response interval manipulation, participants were asked whether they felt the interview and the experience of listening to questions and providing answers had felt normal or awkward. Of the 13 participants in the long interval group, 12 reported that the interactions felt 'normal'; of the 12 participants in the short interval group, 9 reported 'normal'. However, 6 ratings were given with qualifications. Of these 6, 5 participants explained that the questions or the experience of thinking of answers was "unusual" or "not something I had ever thought about", indicating that their ratings were based not on the interaction but instead on the content of the questions. The remaining participant qualified his response by reporting that the interaction felt awkward at first but then seemed normal; he was in the group of participants who received a short interval, which was meant to approximate standard interviewing style. Although the validity of awkwardness ratings collected by the interviewer may be limited, that six participants did give critical responses suggests that participants felt comfortable providing honest responses. Overall, the slow interviewer pacing seemed not to affect participants' view of whether the interview proceeded normally.

Proxy protocol length

Given the inclusion of additional time due to the minimum response interval, participants were asked whether the set of proxy questions felt long or short. Of the 25 participants, 23 reported that the questions felt short. One participant reported that the questions felt "in between" long and short, and the other participant reported that the questions felt long. However, the latter participant qualified her response by noting that the questions felt long because they were difficult to answer. Both participants received the long minimum response interval. Overall, this does not suggest that the slow interview pacing affected participants' subjective feeling of interview length.

Knowledge of others' spending

The quality of recall of others' spending depends on first having knowledge of others' spending. Participants were asked how much they knew about how other people in their household spend money, on a scale ranging from 'A great deal' to 'Not at all' (Table 8). Of the 25 participants, all felt that they were at least somewhat knowledgeable about others' spending. In debriefing,

several participants remarked that they obtain the knowledge through discussions with others, by noticing new purchases, or by being the parent who controls a child's income. Several participants also noted that there are some expenses they just do not know about. For example, one participant said: "I just [only] know he got the tennis shoes - I was with him...I didn't want to say anything [add more expenses] because I don't know." As expected, most participants said they knew only generally the kinds of things that others' spend their money on, but not the amounts they spent exactly.

Table 8. Ratings of knowledge of other household members' spending (n=25)

<i>Rating</i>	<i>n</i>
A great deal	8
Quite a bit	12
Somewhat	5
Very little	0
Not at all	0

CEQ expenditures

Participants were administered selected sections from the standard CEQ. The sections were chosen to be relevant to the information provided by the participant during proxy questioning. A summary of the items and amounts reported is shown in Table 9. The data have been combined across all participants, regardless of whether the section was being probed for proxy reporting (n probed in Table 9).²

Proxy expenditures added after probing

After completing a probed CEQ section, the participant was reminded of the information given during the proxy questions to prompt the participant to think of items not yet reported. As described in Appendix B, the initial round of interviews used a highly standardized and impersonal probe. None of the first 11 participants added expenditures and so the protocols were revised, as described on page 2 and Appendix B.

The revised protocol script elicited at least one additional expense from 5 out of 14 participants. These additional items included:

- Fees for participating in races: \$30, \$58
- Shoes for a visiting child: \$30
- Fees for participating in kayaking: \$350 (\$50 every 2 weeks in the reference period)
- Tennis balls: \$60 (\$15 every month in the reference period)
- Gasoline for mower: \$30
- Jeans: \$30

Overall, the revised protocol successfully elicited additional expenditures. The probes collected extra expenditures that would not have otherwise been recorded. The additional reports yielded

² The only difference between being probed and not being probed that is relevant to initial CEQ reporting is the insertion of a sentence at the start of the CEQ section "When we started, you mentioned that [name] might have had some of these expenses. Please keep them in mind as you answer the questions." An analysis of the effect of this reminder on CEQ reporting was planned but could not be conducted because, as a result of the on-the-fly selection of CEQ sections, there were no sections with an appropriate balance of probed and not probed participants.

by the probes totaled \$588 and averaged \$98 per report. The reports represented between 0.9% and 8% of the reporting participant's total expenditures reported during the CEQ.

Table 9. Summary of CEQ reporting, before additional proxy probing³

	<i>n</i>	<i>n</i> <i>Probed</i>	<i>Count</i> <i>Items*</i>	<i>Items</i> <i>/ n</i>	<i>\$/</i> <i>n</i>	<i>\$/</i> <i>Items</i>
Sect 6: Appliances, household equipment, and other selected items	10	8	75	7.5	\$1,467.2	\$195.6
Sect 8: Home furnishings and related household items	17	3	228	13.4	\$1,451.2	\$108.2
Sect 9: Clothing and clothing services	18	12	437	24.3	\$1,076.6	\$44.3
Sect 12: Vehicle operating expenses	5	5	39	7.8	\$1,356.2	\$173.9
Sect 16: Educational expenses	2	2	7	3.5	\$1,520.0	\$434.3
Sect 17: Subscriptions, memberships, books, and entertainment expenses	20	18	294	14.7	\$1,302.4	\$88.6
Sect 18: Trips and vacations**	12	11	29	2.4	\$2,125.4	\$586.3
Sect 19: Miscellaneous expenses	1	1	5	5.0	\$326.0	\$65.2

*The number of items reported is approximate. Where the data were available, precise figures have been used. In some instances, participants reported monthly expenditures (e.g., spending on jewelry, iTunes music, or trips to the cinema, monthly). In other instances, where participants did not specify quantity, the interviewer used judgment to estimate the quantity.

**For trips and vacations, each trip reported is indicated as an 'item'. Due to a change in protocol to focus on identifying trips rather than enumerating expenses, four participants were not asked to report expenditures and so are omitted from the calculation of \$/n.

The proxy question about household members' hobbies had the most impact, yielding a total of four additional reports. The proxy question about changes in day-to-day routines led to one additional report and the question about ways that others spend their money led to one additional report. Although the conclusions that can be drawn from this information are limited, the impact of the hobbies question indicates that it may be the most effective proxy question of those tested.

Retrospective ratings of proxy question usefulness

To assess whether the proxy questions were useful in raising participants' awareness of how others' spend money, participants were asked whether they would have remembered their household members' purchases even if the interviewer had not reminded them. Respondents were evenly split: 12 participants responded 'yes', 12 responded 'no', and 1 participant was not asked.

³ One participant reported what seemed to be inaccurate, high estimates of amount spent—expenses such as three members of the family going to the movies throughout the reference period was reported as \$1,500 and expenses totaled \$25,369 for the 3-month reference period. When removing this participant, the overall averages for each section remain similar except for reporting in Section 17: spending per participant dropped from \$1,302.4 to \$342.3 and spending per item dropped from \$88.6 to \$25.8.

Another question asked participants to rate how helpful the proxy questions were in reminding them about others' expenses. The results are shown in Table 10. All participants felt that the questions were at least somewhat helpful; one participant was not asked. Participants who did not report additional expenditures when probed still found the proxy protocols helpful. In other words, the initial proxy questions may have caused participants to remember additional expenditures before the directed probing occurred at the end of each CEQ section.

Table 10. Ratings of helpfulness of proxy questions (n=24)

	<i>n</i>
1 Not at all helpful	0
2	0
3	4
4	5
5 Very helpful	15

Considering these results, the introduction could be revised to increase cooperation from respondents and ensure respondent understanding of the purpose; only a few participants understood the purpose of the proxy questions when asked during debriefing and at least one participant thought that he did not need to later report the expenses mentioned during the initial proxy protocol.

Timing of proxy questions

The approximate time spent administering the additional proxy questions was recorded using a CAPI instrument. The total duration of the questions was measured from when the interviewer began reading the introduction for the proxy questions until when the interviewer finished typing the participant's response to the final proxy question. The time does not include the time spent probing during the CEQ reporting sections, as the questions collected expenditure data. The measures of average time spent are shown in Table 11; the median time across all 25 participants was 2m 41s. This median duration for the five proxy questions tested in the current study is similar to the median reported by Book and Edgar (2012) of 2m 15s for four questions.

By design, half of the participants received a longer minimum response interval (a minimum of 10 seconds elapsing before continuing to the next question). When comparing across groups, the participants who received a long minimum response interval spent no longer than the participants who received a short interval (long interval group median = 2m 21s, short interval group median = 2m 52s). When examining mean duration and accounting for a participant in the short interval group whose duration was more than two standard deviations longer than the mean (Table 12), the groups are also no different. Indeed, the results show that the participants in the long interval group on average completed the questions in less time.

Table 11. Average time (min:sec) spent on initial proxy questions

	<i>Median</i>	<i>Mean</i>	<i>SD</i>	<i>Minimum</i>	<i>Maximum</i>
All participants (n=25)	2:40	2:58	1:15	1:33	7:06
Short interval (n=12)	2:52	3:16	1:38	1:33	7:06

Long interval (n=13)	2:21	2:41	0:40	2:00	3:48
-------------------------	------	------	------	------	------

Table 12. Average time (min:sec) spent on initial proxy questions (outlier omitted)

	<i>Median</i>	<i>Mean</i>	<i>SD</i>	<i>Minimum</i>	<i>Maximum</i>
All participants (n=24)	2:31	2:47	0:55	1:33	5:44
Short interval (n=11)	2:45	2:53	1:12	1:33	5:44
Long interval (n=13)	2:21	2:41	0:40	2:00	3:48

Breaking the durations down by total time spent on each of the proxy questions provides insight into why there are no differences between the two timing groups (Table 13). The overall median total time spent on each question by each participant in the short interval group is 14.5s, which is above the minimum 10s mandated for the long interval group.

Table 13. Average time (sec) spent on individual proxy questions

	<i>Median</i>	<i>Mean</i>	<i>SD</i>	<i>Minimum</i>	<i>Maximum</i>
<i>Short Interval</i>					
Other Types	31.2	45.6	52.6	4.2	198.0
Trips	7.1	14.0	17.7	3.5	67.0
Routines	19.6	29.9	37.3	3.1	138.3
Unusual	8.3	10.8	8.9	2.9	34.1
Hobbies	15.6	21.2	16.3	3.6	53.4
<i>Long Interval</i>					
Other Types	24.9	32.8	21.5	12.6	83.9
Trips	12.1	14.8	5.2	11.1	27.7
Routines	12.3	13.4	3.2	10.9	22.3
Unusual	11.8	15.5	6.2	10.7	26.5
Hobbies	13.0	15.8	5.7	10.9	28.0

These findings suggest that slowing interviewer pacing has a smaller effect on overall duration than individual respondent differences.

Discussion

The primary purpose of this study was to refine and standardize the wording of proxy reporting questions and probes for use in production settings. The design of the study built upon the methods and results of Book and Edgar (2012), which developed a protocol for raising

respondents' awareness of how others in the household spend their money. Overall, the final protocols tested in the current study seem to balance sensitivity and burden while still yielding additional expenditures that would not otherwise be captured by standard CEQ interviewing. The final protocols used represent a standardization of the question scripting developed by Book and Edgar (2012) and the analyses presented in this report indicate proposals for refinement of which questions should be asked and which may be considered for omission.

Initially, this study strictly standardized the proxy questions. However, intermediate results indicated that the revisions had decreased the effectiveness of the proxy questions as compared to Book and Edgar (2012) and protocols were modified to increase the level of detail included in the scripts. With the modified scripts, the probes did yield additional reports. However, where the previous study had found over \$3,200 in additional expenditures from 10 out of 18 people, the current study found \$588 from 5 out of 14 people using the modified scripts. Although it is not advised to directly compare the 2012 study, the first half of the current study, and the second half of the current study (at the very least, the 2012 and current study differ methodologically on one major point: different interviewers), a cautious judgment of the overall pattern of results suggests that tailoring a probe and increasing detail improves its effectiveness at eliciting additional expenditures. Additionally, it is noteworthy that the relatively low level of personalization used in the final protocol was successful while also being simple for FRs to implement as it requires only inserting personal information in a standardized probe.

The results from this study can be used to refine future proxy protocols. The questions about how others spend their money, what hobbies or activities others in the household like to spend time on, and any changes in day-to-day routines were the only probes to elicit additional expenses and should be considered for inclusion in future protocols. In contrast, several participants reported that the 'unusual purchases' question felt sensitive while the majority of respondents reported no trips taken by others in their household without them. Given these considerations, these latter two questions may not be effective for a broad range of people and should be considered for exclusion, thus reducing the duration of the set of proxy questions and potentially reducing protocol sensitivity. Given this reduction in length, it may be expected that the protocol would require less than or about two minutes to administer. Additionally, it is recommended that different family situations be considered: for example, a single parent reporting on behalf of a household of young children should not be asked questions about how their children spend money. To address this, the protocol could only be administered to CUs of more than two people and only include questions about people who control their own purchases (e.g., older than 10 years of age and not in poor health). This is expected to raise the effectiveness and efficiency of the protocol while also reducing the likelihood that respondents would view the questions as irrelevant.

A secondary goal of this study was to understand how interviewer pacing for open-ended questions affects response quality. The results supported the hypothesis that longer minimum response intervals were associated with an increased number of reports. Notably, the minimum response interval also enabled at least two participants to change their answers. These benefits were gained without a costly increase in subjective interview duration or decrease in interaction quality.

There are several important limitations to this work that should be considered when interpreting these results. The study is based on a small non-representative sample and this challenge was further limited by the proportion of participants who lived in households with only young children that do not spend money. Additionally, it was not possible to check the validity of expenses reported and so reports of higher numbers of items and amounts were assumed to be more

accurate. And, as with many interviews, the scripting throughout the interview (proxy questions and CEQ) was not always followed precisely, which could lead to unaccounted for differences between groups.

The results from the current study and Book and Edgar (2012) indicate that proxy protocols warrant further research and testing. Although reliable estimates of effect size are not possible to make from this study's small sample size, the results indicate that the protocols led to increased reporting. However, given that the expected variance in respondent reporting is high, it is not recommended that future research hypotheses be based solely on comparisons of averages between groups to test different protocols. It is recommended that future research investigate the proxy reporting question 'incrementally'—building from the simplest protocol and adding manipulations to collect the effect of incrementally richer protocols as compared to baseline. If an incremental change in protocol leads to an additional expense reported, we can be fairly confident that the additional report is due to the incremental change. An additional manipulation to test the effectiveness of the proxy questions alone without probes may also be of value. The current design, in which all participants receive the initial proxy questions, can only provide a conservative estimate of the true effect of the proxy protocols as compared to the current production procedures. Given the lower demands on interviewer ability of such an approach (i.e., no need to prompt respondents throughout the CEQ), it may prove successful. Although these approaches confound participant time spent on the question, they seem to provide the most straightforward comparisons between multiple protocol designs.

Recommendations

Recommendation 1: Target only multi-person CUs with at least one other CU member who controls their own spending. Heuristics such as a minimum age of 10 years and no reports of poor health could be used to aid the FR in judging whether to include a CU member though future research should test this question (cf., Cate-Schaeffer, 2010; Edgar, 2011).

Recommendation 2: Use rich and personalized probes that include specific details and events to trigger memory.

Recommendation 3: Use the proxy questions about how others spend their money, what hobbies or activities others in the household like to spend time on, and any changes in day-to-day routines. Do not use the questions about unusual purchases and trips.

Recommendation 4: Encourage interviewers to allow respondents more time to reflect on open-ended questions and give thoughtful answers.

References

- Blair, E. & Burton, S. (1987). Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research*, 14, 280-288.
- Book, T. & Edgar, J. (2012). Proxy Reporting Lab Study Report. Bureau of Labor Statistics internal report. Washington, DC: Bureau of Labor Statistics.
- Bureau of Labor Statistics (2010). Proxy Reporting in the Consumer Expenditure Surveys Program. Bureau of Labor Statistics internal report. Washington, DC: Bureau of Labor Statistics.
- Burton, S. & Blair, E. (1991). Task conditions, response formulation processes, and response accuracy for behavioral frequency questions in surveys. *Public Opinion Quarterly*, 55, 50 -79.
- Cate-Schaeffer, N. (2010). Issues in proxy reporting in the Consumer Expenditure Survey. Paper presented at CE Methods Workshop. Washington, DC: Bureau of Labor Statistics.
- Edgar, J. (2011). 2010 CE methods workshop summary. Bureau of Labor Statistics internal report. Washington, DC: Bureau of Labor Statistics.
- Kojetin, B. & Jerstad, S. (1997). The quality of proxy reports on the Consumer Expenditure Survey. Bureau of Labor Statistics internal report. Washington, DC: Bureau of Labor Statistics.
- Mathiowetz, N. (1987). Questionnaire Design Advisory Conference. Bureau of Labor Statistics internal report. Washington, DC: Bureau of Labor Statistics.
- Reiser, B. J., Black, J. B., & Abelson, R. P. (1985). Knowledge structures in the organization and retrieval of autobiographical memories. *Cognitive Psychology*, 17, 89-137.
- Schwarz, N. & Wellens, T. (1997). Cognitive Dynamics of Proxy Responding: The Diverging Perspectives of Actors and Observers *Journal of Official Statistics*, 13, 159–179.

Appendix A: Proxy question scripting and protocol

You said that [name1, name2, and name3] live in your household. I'd like to spend a few minutes talking about ways they spend their money. This will help us move faster later in the interview, and help you remember things that you might otherwise forget since they're not your expenses. It's best if you take time to really think about each answer for each person. I'll give you a moment to form your answer to each question.

To start, I want to ask about expenses that the other people in your household have that you may not. For example, one person might buy all the clothing, while another handles all the vehicle expenses.

Are there some types of things that [name1] spends money on that you don't?

Interviewer starts the timer; the interviewer does not move on until the minimum response interval has passed.

When the participant starts speaking, the interviewer records the participant's response.

When the participant finishes responding, the interviewer asks about the next person in the household.

Are there some types of things that [name2] spends money on that you don't?

The interviewer repeats the question for all members of the household, checks that the minimum response interval has passed, and then stops the timer and moves on to the next question.

For the remaining four questions, the interviewer starts the timer after reading the question and does not move on until the minimum response interval has passed. After the participant finishes responding, the interviewer moves on to the next question.

Thinking about the last three months since May 1st, has [name1], [name2], or [name3] taken any vacations or trips without you?

Thinking about the last three months since May 1st, has [name1], [name2], or [name3] made any changes to their normal day-to-day routine that might have changed what they did or did not spend money on?

Thinking about the last three months since May 1st, has [name1], [name2], or [name3] bought anything unusual or out of the ordinary?

Does [name1], [name2], or [name3] have any hobbies or activities that they like to spend time on?

Appendix B: Modifications made to probe scripting compared to original scripting

	Original Scripting (n=11)	Modified Scripting (n=14)
Things others spend money on that you don't	You mentioned that [name] spends money on [reported item]. Are there any other expenses that you want to add related to that?	You mentioned that [name] spends money on [reported item]. Did [name] have any other [item type] expenses that we haven't talked about yet?
Trips	You mentioned that [name] took a trip recently. Are there any other expenses that you want to add related to that trip?	You mentioned that [name] took a trip recently. Did [name] have any other [item type] expenses that we haven't talked about yet?
Changes to day-to-day routine	You mentioned that [name] made a change to their day-to-day routine. Are there any other expenses that you want to add related to that?	You mentioned that [name] [change to their day-to-day routine]. Did [name] have any other [item type] expenses that we haven't talked about yet?
Unusual purchases	You mentioned that [name] bought something unusual or out of the ordinary. Are there any other expenses that you want to add related to the unusual item?	You mentioned that [name] bought [something unusual or out of the ordinary]. Did [name] have any other [item type] expenses that we haven't talked about yet?
Hobbies	You mentioned that [name] has a hobby or activity they like to do -- [reported hobby or activity]. Are there any other expenses that you want to add related to that?	You mentioned that [name] has a hobby or activity they like to do -- [reported hobby or activity]. Did [name] have any other [item type] expenses that we haven't talked about yet?

Appendix C: Debriefing questions

Debriefing after initial proxy questions

- What did you think of the questions we just asked you about what the other people in your household spend money on, their trips, routines, unusual purchases, and hobbies?
- Why do you think we asked those questions?
- Do you think you were able to answer them accurately?
- How much do you know about how other people in your household spend their money? (A GREAT DEAL/QUITE A BIT/SOMEWHAT/VERY LITTLE/NOT AT ALL)
- Do you know the types of things they spend their money on generally? (YES/NO)
- Do you know the amounts of money they spend, generally? (YES/NO)
- Do you know exactly? (YES/NO)
- How long would you say that first set of questions was? Long or short? (LONG/SHORT)
- How sensitive, or personal, would you say those questions were? On a scale from 1 to 5, where 1 is 'not at all sensitive' and 5 is 'very sensitive'. (1 2 3 4 5)
- What do you think other people would say, on a scale from 1 to 5? (1 2 3 4 5)
- How did you feel about the experience of listening to the questions and giving your answers? Did it seem normal or awkward? (NORMAL/AWKWARD)

Debriefing after CEQ reporting sections

- Based on your answers to those questions, when we got to the sections about [insert expenditure section with proxy probes], I reminded you about the purchases made by other members in your household. If I had not instructed you to remember their purchases, do you think you would have? (YES/NO)
- How helpful were the questions at the start about what others spend their money on in reminding you about others' expenses, on a scale from 1 to 5, where 1 is 'not at all helpful' and 5 is 'very helpful'? (1 2 3 4 5)
- Do you think you included all the expenses that people in your household had? (YES/NO)
- I did not give extra instructions for [insert expenditure category asked without proxy probe OR clothing]. I just said [insert expenditure category transition]. Let's go back and review those questions now. I asked you about this list of items. Do you think you included all the expenses that people in your household had?
Show information booklet page (YES/NO)
- How helpful would it have been to have additional questions to jog your memory during this section, on a scale from 1 to 5, where 1 is 'not at all helpful' and 5 is 'very helpful'? (1 2 3 4 5)
- We're interested in helping people tell us as much as they can about expenses for all the members of their household. Do you have any suggestions for how we could have helped you report everyone's expenses more accurately?