

ABSTRACT

Title of dissertation: CLARIFYING SURVEY QUESTIONS

Cleo Redline
Doctor of Philosophy, 2011

Dissertation directed by: Professor Roger Tourangeau
Joint Program in Survey Methodology

Although comprehension is critical to the survey response process, much about it remains unknown. Research has shown that concepts can be clarified through the use of definitions, instructions or examples, but respondents do not necessarily attend to these clarifications. This dissertation presents the results of three experiments designed to investigate where and how to present clarifying information most effectively. In the first experiment, eight study questions, modeled after questions in major federal surveys, were administered as part of a Web survey. The results suggest that clarification improves comprehension of the questions. There is some evidence from that initial experiment that respondents anticipate the end of a question and are more likely to ignore clarification that comes after the question than before it. However, there is considerable evidence to suggest that clarifications are most effective when they are incorporated into a series of questions. A second experiment was conducted in both a Web and Interactive Voice Response (IVR) survey. IVR was chosen because it controlled for the effects of interviewers. The results of this experiment suggest that readers appear no more capable of comprehending

complex clarification than listeners. In both channels, instructions were least likely to be followed when they were presented after the question, more likely to be followed when they were placed before the question, and most likely to be followed when they were incorporated into a series of questions. Finally, in a third experiment, five variables were varied to examine the use of examples in survey questions. Broad categories elicited higher reports than narrow categories and frequently consumed examples elicited higher reports than infrequently consumed examples. The implication of this final study is that the choice of categories and examples require careful consideration, as this choice will influence respondents' answers, but it does not seem to matter where and how a short list of examples are presented.

Clarifying Survey Questions

by

Cleo D. Redline

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:
Professor Roger Tourangeau, Chair
Professor Katharine G. Abraham
Professor Frederick G. Conrad
Professor Kent L. Norman
Professor Norbert Schwarz

© Copyright by
Cleo Redline
2011

Dedication

In loving memory of my grandparents, Elizabeth McHenry and
Elsie and Richard Redline.

Acknowledgments

I have a great deal to be thankful for and many people to thank. I am greatly indebted to my dissertation advisor, Roger Tourangeau, for his unwavering mentorship throughout the process of producing this dissertation, despite the many obstacles and challenges that presented themselves along the way. I know that I am a better researcher as a result of working with him. As anyone knows who has worked with him, or read his work, his intellect is great (he is extremely perspicacious) and his standards are high (his exacting thoroughness is unparalleled); consequently, I had no choice but to grow in his presence. I am grateful for having been given the opportunity to do so.

I would also like to thank my advisory committee members – Katharine Abraham, Fred Conrad, Kent Norman, and Norbert Schwarz – for giving of their time to review and comment on my work. I have no doubt that this dissertation is a better product as a result of their contributions. I would also like to thank Bob Groves for encouraging me when my ideas for this dissertation were still emerging, and I would like to thank the all of the professors and my fellow students in the Joint Program in Survey Methodology (JPSM) for their inspiration along the way, their words of encouragement, and their friendship.

The research presented here was supported by a number of agency grants and people to whom I owe an enormous debt of gratitude. The first experiment was supported by a National Science Foundation (NSF) Major Instrumentation Grant to Jon Krosnick. Jon was kind enough to allow me space on one of his Web panels,

as a courtesy to Roger. I am grateful to them both. The second experiment was supported by an U.S. Bureau of the Census Doctoral Dissertation Fellowship to me (with Roger as my advisor) and a NSF Doctoral Dissertation Research Improvement Grant to Stanley Presser and me. I will be eternally grateful to the Census Bureau for having awarded me their fellowship – I fear this dissertation would not have materialized without it – and to Stanley for serving as principal investigator on the NSF grant. In the case of the third experiment, I am beholden to Roger, Mick Couper, Fred Conrad, and Reg Baker for generously allowing me space on one of their Web panels, which was supported by a National Institutes of Health grant to them.

I would also like to acknowledge several people who were instrumental in the conduct of the experiments presented herein. Allison Ackermann and Dean Williams, both of Abt SRBI, managed the first and second experiments, respectively. They were consummate professionals, offering advice and responding to my every request without hesitation. I would also like to thank Joanne Mechling from Market Strategies International for her contribution to the implementation of the third experiment. In addition, I would like to thank the many anonymous respondents who answered my questions. Finally, the research presented here would not have been possible without the key assistance of Duane Gilbert, Rupa Jethwa Eapen, and Sarah Gebremicael from JPSM.

Being a part of Roger's research group these past few years has been a great learning experience. It gave me the opportunity to work with some of the finest minds in survey research – Fred Conrad, Mick Couper, and Reg Baker – and to

stretch my abilities. It also gave me the opportunity to forge new, and I hope, lasting friendships.

I would be remiss if I did not mention the many wonderful people I had the great privilege to get to know and to work with while I was at the Bureau of the Census. The Center for Survey Methods Research was an exciting and stimulating place for a budding researcher such as myself to work and I was extremely fortunate to have been given the chance to work with two people I greatly admire – Don Dillman and Betsy Martin.

Finally, I would like to thank my beloved family, who saw little of me and who learned not to expect much from me during my time at Maryland. I am especially grateful to my mother, Zoe Kritzer, who instilled in me a love of learning, who always had words of encouragement and who always seem to think things would work out for the best (even when I was sure they would not). This dissertation is the culmination of her faith and her values.

This last position is reserved for the two people I hold most dear to me and to whom I owe the most: my husband, Robert Creecy, to whom I turned for countless (and I do mean countless) advice, support, and comfort during the production of this dissertation and my daughter, my precious, precious daughter, Tori YuXiang Creecy, who I fear sacrificed the most while I carried out this endeavor. I am eternally grateful to them both for the love they bestowed upon me and the sacrifices they made, and can only hope that I will be able to make it up to them in the years ahead.

Table of Contents

List of Tables	viii
List of Figures	x
1 Comprehension Problems in Surveys	1
1.1 Sources of Comprehension Problems Related to Clarification	3
1.1.1 Lexical Ambiguity	7
1.1.2 Vagueness and Conceptual Variability	8
1.1.3 Complicated Mappings	9
1.2 Solutions to Comprehension Problems	11
1.2.1 Interviewer Interventions	11
1.2.2 Definitions	12
1.2.3 Instructions	12
1.2.4 Examples	13
1.3 Obstacles to the Use of Clarification	17
1.3.1 Presumption of Interpretability	17
1.3.2 Conversational Order	20
1.3.3 Other Factors Affecting the Use of Clarification	23
1.4 Sensory Channel Effects	26
1.4.1 Channel and the Presumption of Interpretability	27
1.4.2 Channel and Order	28
1.4.3 Channel and Other Factors	29
1.5 Summary of the Literature	31
2 Clarifying Instructions in a Web Survey	35
2.1 Introduction	35
2.2 Methods	39
2.2.1 Study Items	39
2.2.2 Experimental Conditions	41
2.2.3 Accessing the Effectiveness of the Clarifying Instructions	42
2.2.4 Data Collection and Sample	44
2.3 Results	45
2.3.1 Responses to the Study Items	46
2.3.2 Responses to the Follow-up Questions	53
2.3.3 Response Times	57
2.4 Discussion	60
2.5 Conclusions and Future Research	65
3 Crossing Clarifying Methods and Sensory Channels	67
3.1 Introduction	67
3.2 Methods	69
3.2.1 Sampling and Data Collection	70
3.2.1.1 IVR Cases	71

3.2.1.2	Web Cases	72
3.2.2	Study Questions and Clarifying Methods	73
3.3	Results	77
3.3.1	Effects of Nonresponse	77
3.3.2	Responses to the Study Items	77
3.3.3	Zero Responses	88
3.3.4	Responses to the Follow-up Questions	88
3.3.5	Response Times	91
3.4	Discussion	93
3.5	Conclusions and Future Studies	98
4	Varying Examples and Their Presentation in a Web Survey	101
4.1	Introduction	101
4.2	Methods	105
4.2.1	Study Questions and Experimental Conditions	105
4.2.2	Data Collection and Sample	108
4.3	Results	109
4.3.1	Mean Number of Servings	109
4.3.2	Response Times	111
4.4	Discussion	114
5	Conclusions	116
5.1	Summary	116
5.2	Limitations	123
5.3	Future Research	125
	Appendices	127
A	Questionnaires	127
A.1	Experiment 1	127
A.2	Experiment 2	147
A.3	Experiment 3	174
	Bibliography	198

List of Tables

2.1	Examples of the No Clarifying Instruction and Alternative Clarifying Instruction Methods	43
2.2	Mean Response (and Sample Sizes) by Main Condition and Item	49
2.3	Mean Response (and Sample Sizes) by Order of Presenting Clarifying Instructions and Item	51
2.4	Percentage of “Valid” Responses by Method of Presenting Clarifying Instructions for the Follow-up Items	56
2.5	Mean Response Time in Seconds (and Sample Sizes) by Order of Presenting Clarifying Instructions and Item	59
3.1	Examples of Alternative Clarifying Methods	75
3.2	Percentage of Respondents Who Accessed Either Computer System (and Sample Sizes) by Outcome, Mode of Data Collection, and Method of Clarification	78
3.3	Percentage of Respondents (and Sample Sizes) by Demographic Characteristics or Reported Behaviors and Mode of Data Collection	79
3.4	Percentage of Out-of-Range Values and Percentage of Item Missing (and Sample Sizes) by Mode of Data Collection and Method of Presenting Clarifying Instructions	81
3.5	Mean Response (and Sample Sizes) by Method of Presenting Clarifying Instructions, Mode of Data Collection, and Item	82
3.6	Individual Statistics (ANOVA) for Main and Interaction Effects, by Item	83
3.7	Mean Percentage of Zeros Over All Items (and Sample Sizes) by Method of Presenting Clarifying Instructions and Mode of Data Collection	89
3.8	Percentage of “Valid” Responses (and Sample Sizes) by Order of Presenting Clarifying Instructions, Mode of Data Collection, and Follow-up Item	90
3.9	Mean Response Time in Seconds (and Sample Sizes) by Order of Presenting Clarifying Instructions, Mode of Data Collection, and Item	92

4.1	Examples by Breadth and Frequency of Consumption	106
4.2	Mean Number of Servings (and Sample Sizes) by Item and Experimental Condition	110
4.3	Mean Response Time in Seconds (and Sample Sizes) by Item and Experimental Condition	112
5.1	Means of the Before and Multiple Question Methods Relative to the After Method for Experiments 1, 2, and 3	123

List of Figures

2.1	Mean Response to Items as a Function of the Main Conditions	48
2.2	Standardized Mean Response to Items as a Function of the Order of Presenting Clarifying Instructions	52
3.1	Standardized Mean Response Across Items as a Function of Clarifying Method and Channel for All Items (top panel), Items Involving Low Counts (middle panel) and High Counts (bottom panel)	86
4.1	The Order, Font, and Format of the Examples for One Item	107

“But I would like to explore a lesser-known debate triggered by 9/11. Exactly how many events took place in New York on that morning in September?... The 9/11 cardinality debate is not about the facts, that is, the physical events and human actions that took place that day... But the construal of those facts: how the intricate swirl of matter in space ought to be conceptualized by human minds. As we shall see, the categories in this dispute permeate the meanings of words in our language because they permeate the way we represent reality in our heads. ” —Steven Pinker, The Stuff of Thought: Language as a Window into Human Nature, 2007

Chapter 1

Comprehension Problems in Surveys

Although comprehension is well recognized as a critical component of the survey response process (e.g., Tourangeau, Rips, & Rasinski, 2000, p. 9), much about it remains unknown. Linguists and psychologists believe that there is often a complex relationship between the meaning of words and the way in which we categorize information (e.g., Lakoff, 1987; Ruhl, 1989; Smith, 1995). Studies have observed comprehension problems, such as lexical ambiguity or vagueness, in respondents’ attempts to answer survey questions (e.g., Suessbrick, Schober, & Conrad, 2000). Clarifying the question’s intended meaning when the question uses ambiguous or vague terms may improve response accuracy. For example, clarifying what should be and should not be counted as “furniture” has been shown to aid respondents’

understanding of what to report in a question about furniture purchases (Conrad & Schober, 2000). Although the benefit of providing this clarifying information appears clear, where and how to present it, such that respondents view processing it as an essential part of their task, is less clear. Thus, the first aim of this research is to investigate how to present clarifying information so that respondents will recognize it as essential to answering questions correctly.

On theoretical grounds, whether the question and clarifying information is presented aurally or visually may influence the degree to which the question is understood, especially if it is complex. In theory, readers (provided they read sufficiently well) should be able to understand complex clarifying information better than listeners (Clark & Brennan, 1991; Just & Carpenter, 1980; Osada, 2004; Rayner & Clifton, 2009; Schwarz, Strack, Hippler, & Bishop, 1991a). However, the channel in which the information is presented may interact with how the clarification is presented to influence respondents' answers, but to date, there is little empirical evidence to confirm these hypotheses. Thus, the second aim of this research is to investigate whether the channel of presentation influences respondents' comprehension of clarifying information in surveys, and whether the channel interacts with whether the clarifying information is presented before or after the question itself.

In the absence of any clarification, respondents may erroneously expand or restrict their interpretation of a survey question because the question evokes examples that differ from the survey's intentions. For instance, a question about furniture may erroneously bring floor lamps to mind. At the root of these problems appears to be the fact that words can evoke different meanings due to differences in the

way we categorize information (Lakoff, 1987). As Suessbrick, Schober, & Conrad (2000) point out, some respondents may categorize floor lamps as furniture, while others may not. Still others may do so under some circumstances, but not other circumstances. Questions sometimes provide examples, either in part or exclusively, to remedy this situation. However, it is not clear what examples to present, how to present them, and what the effects of providing examples are. Thus, the third aim of this research is to investigate the effects of presenting different types of examples and different methods of presenting them.

1.1 Sources of Comprehension Problems Related to Clarification

The asking and answering of questions is at the heart of the survey process. A number of cognitive models have been proposed to describe this process (Cannell, Miller, & Oksenberg, 1981; Tourangeau, 1984). These models agree that a key component is comprehension of the question. Comprehension encompasses such processes as attending to the question and accompanying instructions, assigning a meaning to the surface form of the question, and inferring a question's point (Tourangeau, Rips, & Rasinski, 2000, p. 9). The goal is that every respondent answering a question should understand it in a consistent way and in a way that is consistent with what the researcher meant (Fowler, 1995, p. 2).

Respondents' understanding of questions are not only influenced by the semantic meaning of words (the meanings typically attached to the words themselves), but by the pragmatic meaning of an utterance (the speaker's intended meaning) (e.g.,

Schwarz, 1996, p. 7). Accordingly, respondents assume that a survey is governed by the same maxims that govern conversations (Grice, 1975). These maxims — that speakers will be truthful, relevant, informative and clear — have many implications for respondents’ understanding of questions. If a question is ambiguous, and the survey does not provide additional clarification, respondents may turn to the question’s context to determine its meaning (e.g., Schwarz, 1996; Tourangeau & Rasinski, 1988). False implicatures occur when respondents extract unintended meanings from questions or response categories because they presume the survey is operating under the Gricean maxims. For example, under the assumption that a survey would not ask the same question twice (because this violates the maxim to be informative), respondents who have just been asked a question about their marriage may incorrectly assume that a subsequent question about life in general pertains to other aspects of their life (Schwarz, Strack, & Mai, 1991b; Tourangeau, Rasinski, & Bradburn, 1991). Such inferences may be especially problematic in standardized surveys, since respondents are unable to confirm the intended meaning of the words with the interviewer (Clark & Brennan, 1991; Clark & Schober, 1992). Thus, to infer the intended meaning of a question in the absence of further clarifying information, respondents may attend to a wide range of cues – format, question context, the range of the response alternatives, information about the researchers’ affiliation, the survey sponsor, and the visual features of the questionnaire (e.g., Schwarz, Grayson, & Knauper, 1998; Schwarz & Oyserman, 2001; Tourangeau, Couper, & Conrad, 2004).

One of the most frequently cited studies to suggest that question wording is

an important problem in surveys is Belson's (1981) study. He detailed respondents' misunderstandings of 29 questions. He found, for example, that 37% of the study respondents misunderstood the phrase "days of the week" in the question "How many days of the week, do you usually watch television? I mean weekdays and Saturdays and Sundays, of course, and daytime viewing as well as evening viewing?". Smaller percentages also misunderstood the terms "you," "usually," and "watch television."

These misunderstandings were uncovered in a second, intensive interview conducted a day after the original interview. The reinterview included probes such as "When you were asked that question yesterday, exactly what did you think it meant?". There are potentially two problems with this approach. For one, the delay between the original and reinterview means information available to the respondent at the time of the reinterview may differ from the information available during the original interview. Of even greater concern, perhaps, is whether the relevant information about the question-answering process can even be articulated by the respondent (Forsyth & Lessler, 1991). The more general a word's meaning, the harder it is to define (Ruhl, 1989); asking respondents to express a question in their own words confounds misunderstanding with their ability to recast the question in new words (Foddy, 1996). Respondents may know what a question means, yet be unable to express this understanding. In line with this, an empirical evaluation of probing methods revealed that paraphrasing was not as productive as other methods (Foddy, 1998). Alternatively, the use of paraphrasing could lead respondents to manufacture paraphrases that do not reflect their actual understanding. Thus, it is

unclear the extent to which misunderstandings such as those uncovered by Belson (1981) are serious, that is, express genuine mismatches between the respondent's understanding of key concepts and the researcher's.

Tourangeau, Rips, & Rasinski (2000, pp. 34-61) describe seven sources of comprehension errors that may lead to mismatches between the respondent's understanding of concepts and the researcher's (see also Graesser, Cai, Louwerse, & Daniel, 2006; Graesser, Kennedy, Wiemer-Hastings, & Ottati, 1999). Two of the seven sources of comprehension errors, *lexical ambiguity* and *vague concepts*, appear to correspond with the problems Belson (1981) identified. In addition, Conrad and Schober and their colleagues have shown that comprehension errors occur when respondents have to map their situations onto survey concepts in a *complicated* rather than *straightforward* way (e.g., Conrad & Schober, 2000; Conrad, Schober, & Coiner, 2007; Schober & Conrad, 1997; Schober & Bloom, 2004).

At the root of these problems appears to be the fact that words can evoke different meanings due to differences in the way in which we can categorize information (Lakoff, 1987). To categorize information is to group objects that belong together. Our semantic memory allows us to combine similar objects into a single concept, but deciding which objects are similar and belong together is a complex subject of much debate and theorizing (Smith, 1995). Consequently, mismatches between a respondent's categorization of objects and the survey's intended categorization of objects produces comprehension error.

1.1.1 Lexical Ambiguity

Lexical ambiguity occurs when a word has more than one meaning, and the context in which the word is used does not make clear (at least immediately) which meaning is intended. This can occur in a number of ways (Lakoff, 1987). In homonymy, a single word (such as the word “bank”) can have several unrelated meanings. “Bank” can mean a place for money or a place along the river. “Ball” can mean a round object used in games or a lavish formal dance.

In polysemy, words are thought to have different, but related meanings (Klein & Murphy, 2001; Murphy, 1997; Nunberg, 1979). According to Pinker (2007, pp. 110-11), polysemy is everywhere: “Window” can refer to a pane of glass or an opening. “Chicken” can refer to a kind of animal or a kind of meat. “Newspaper” can refer to an organization or an object. The word “child” can mean “any young person” or it can mean “one’s offspring, regardless of their age”. Billiett, Looseveldt & Waterplas (as cited in Sudman, Bradburn, & Schwarz, 1996, p. 61) provided an example in which respondents offered numbers from twenty to thirty in response to the question, “How many children do you have?” It turned out that the respondents were teachers, who interpreted this question as referring to the “young people” in their classes rather than their personal “offspring.” Thus, polysemy involves cases in which there is one word or phrase with a family of different but related senses that exhibit cognitive organization (Lakoff, 1987).

1.1.2 Vagueness and Conceptual Variability

Vague concepts are another potential source of comprehension difficulties in survey questions. Vague concepts have unclear boundaries (Pinkal, 1995). Take, for example, the question “Do you have a physical, mental, or other health condition which limits the kind or amount of work you can do?” “Limit” lies along a continuum from severely limiting to ever-so-slightly limiting. To what degree of limitation is this question referring? Or another example, “Have you smoked at least 100 cigarettes in you entire life?” Suessbrick, Schober, & Conrad (2000) showed that the concept of “smoking cigarettes” is vague. It can mean “only cigarettes you finished,” or it can mean cigarettes “you finished or partly smoked,” or it can mean “even just one puff.”

Some linguists argue that polysemy is not so common (e.g., Ruhl, 1989). Instead, they argue, words are monosemic, with one single, but highly abstract meaning. Monosemic words allow conceptual differentiation, although this organization appears to be variations or gradations along a scale. An example is our ability to conceive of several types of dogs or several shades of red, but this does not require us to give the word “dog” or “red” several distinct meanings. In this case, dog is an example of an abstract-concrete relationship, with “dog” more abstract than “terrier” and animal more abstract than “dog .”¹

¹Much earlier, Kant (1781) used a similar example in his attempts to describe the notion of schemata: “The concept of a dog signifies a rule in accordance with which my imagination can specify the shape of a four-footed animal in general, without being restricted to any single particular shape that experience offers me...”

Whether a word is ambiguous or merely vague is unequivocal in some cases. For example, it is easy to identify a word, such as “ball,” as ambiguous between two readings of the word (round object versus formal dance). And it is easy to identify the word ‘small’ as vague, for where along the continuum does small become not small? Some words, however, are both ambiguous and vague. The word “child” can mean any young person or it can mean one’s offspring. When it is used to mean young person, the word “child” is also vague because its boundaries are not clear or sharply outlined. Many other words behave similarly. “Fast,” for instance, can mean quick or fixed. When it means quick, it is vague, as it specifies a range or continuum, the boundaries of which are unclear.

1.1.3 Complicated Mappings

A finding that has clearly surfaced is that comprehension errors tend to occur when respondents’ situations map onto the survey concepts in a complicated way (e.g., Conrad, Schober, & Coiner, 2007). For example, a nuclear family made up of a father, mother, and two children maps onto a question about the number of people living in the house in a straightforward way. A family who has a son or daughter living in a dormitory while attending college during the school year, but home the remainder of the year, maps onto the question in a more complicated way.

A series of studies by Conrad and Schober (and their colleagues) have shown that respondents whose situations map onto questions in a straightforward way tend to answer questions very accurately, with or without further clarification (Conrad &

Schober, 2000; Conrad, Schober, & Coiner, 2007; Schober & Conrad, 1997; Schober & Bloom, 2004). It is respondents whose situations map onto the questions in complicated ways who have difficulty answering questions correctly unless they get further clarification. The results of Tourangeau, Conrad, Arens, Fricker, Lee, & Smith (2006) are consistent with this. They found that participants were better at classifying vignettes that closely matched a definition (central instances) than ones that only partly matched it (peripheral instances) (see also Gerber, Wellens, & Keeley, 1996). Central instances have characteristics that respondents see as fitting the definitions of a category; for example, most respondents tend to recognize that beef is meat. Peripheral instances fit less well – for example, it is less clear whether liver is a meat.

The implication of this research is that whether a respondent requires further clarification or not depends in part on his or her circumstances. If a respondent's situation is simple or can be described as one of the central instances of a concept, then chances are he or she will not need much by way of explanation. If a respondent's situation is complicated or can be described as peripheral, further explanation may be needed. If questionnaire designers knew in advance whether a respondent mapped onto a question in a simple or complicated way, they would know in advance whether respondents need additional clarification, and could act accordingly. The problem is questionnaire designers do not know this in advance. Likewise, if respondents understood that they mapped onto a question in a complicated way, they might understand that they need additional clarification and ask for it. The problem is that respondents do not realize this either. As a result, they do not

necessarily perceive a need for clarifying information or a need to expend the effort to obtain clarifying information.

1.2 Solutions to Comprehension Problems

Several methods have been used to reduce comprehension errors in survey settings.

1.2.1 Interviewer Interventions

One method to reduce comprehension errors in surveys is to allow interviewers to provide clarifying information. This method – conversational interviewing – entails providing interviewers with definitions for ambiguous or vague terms administered in the questions and allowing them to impart this information to respondents as they deem necessary. Two studies by Schober and Conrad showed that it reduced measurement error in comparison to the traditional method of not allowing interviewers to provide clarification (e.g., Conrad & Schober, 2000; Schober & Conrad, 1997).

Extending this to Web surveys, Conrad, Schober, & Coiner (2007) found a relationship between various clarification methods and accuracy. For instance, when the user was in control of obtaining clarification by clicking, accuracy was greater than when no clarification was available at all. However, accuracy was better still when the computer system volunteered clarification based on general guidelines regarding the length of time needed to answer. Accuracy was even better yet when the

computer system provided clarification by taking the respondents' age into account. However, accuracy was best of all when respondents always received definitions along with the question.

1.2.2 Definitions

In the conversational method, interviewers provide respondents with definitions for ambiguous or vague terms. Close to half the time interviewers using the conversational method of interviewing simply presented parts of definitions immediately after they read the questions (73 of the 165 cases) (Conrad & Schober, 2000). Respondents comprehended the questions more accurately when they received clarification than when they did not and when interviewers were trained to initiate this clarification rather than relying on respondents to ask for it (Schober, Conrad, & Fricker, 2004). Similarly, accuracy was the highest when definitions were always presented along with questions in a Web survey (Conrad, Schober, & Coiner, 2007) and when questions were rewritten to include clarification (Fowler, 1992). Other studies show that the more respondents read definitions, the more the definition seems to affect their answers (Galesic, Tourangeau, Couper, & Conrad, 2008; Tourangeau, Conrad, Arens, Fricker, Lee, & Smith, 2006).

1.2.3 Instructions

Another type of clarification involves giving the respondent instructions. Researchers have long made the point that respondents do not read and follow in-

structions in self-administered surveys (Frohlich, 1986; Jansen & Steehouder, 1992; Jenkins & Dillman, 1997). Direct evidence for this comes from the fact that questionnaires with branching instructions have higher item nonresponse rates than surveys without such instructions (Messmer & Seymour, 1982.; Turner, Lessler, Hubbard, & Witt, 1992).

Several studies have examined methods for improving branching instructions (Redline & Dillman, 2002; Redline, Dillman, Dajani, & Scaggs, 2003). Alterations which made the branching instructions more visible (such as increasing the size of the instruction or putting it in bold type) attracted respondents' attention. Features that aided in the repair of navigational errors (such as an arrow pointing to a parenthetical phrase) appeared to improve respondents' performance, reducing both errors of omission (that is, respondents skipping questions they were supposed to answer) and errors of commission (that is, respondents answering questions they were not supposed to answer).

1.2.4 Examples

Another way to clarify the meaning of a question is to provide examples, that is, subcategories that are meant to illustrate the category being asked about. The fact that "furniture" includes such examples as "chairs" and "footstools," but does not include such examples as "floor lamps" has been shown to help respondents understand the meaning of the word "furniture" (e.g., Conrad & Schober, 2000; Schober & Conrad, 1997). This may be because semantic categories are

structured in memory in terms of different levels: “furniture” is an example of a superordinate-level category; “chair” is a more basic-level category; and “desk chair” is a subordinate-level category (Biederman, Subramaniam, Bar, Kalocsai, & Fisher, 1999; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). This notion that semantic memory is structured hierarchically appears to share much in common with the linguistic notions about the cognitive organization of words (Lakoff, 1987; Ruhl, 1989).

In general, basic-level objects, such as “chair,” appear to provide an optimal amount of distinguishing information without overwhelming people with too much detail. As a result, people prefer to use basic-level names (Biederman et al., 1999; Rosch et al., 1976). However, linguists believe that words may differ in abstraction in different domains, for different purposes, or with different people. For example, experts prefer to use subordinate categories to identify objects; birdwatchers prefer to say “warbler” rather than “bird” (Johnson & Mervis, 1997).

Survey respondents may expand or restrict the meaning of concepts when a superordinate-level (or abstract) concept evokes different examples from them from those that the researchers intended. For instance, when not explicitly told what to include and exclude as “furniture,” some respondents appear to expand the meaning to include floor lamps. But when explicitly told to exclude these instances, respondents appear to correctly restrict their interpretation (e.g., Conrad & Schober, 2000; Schober & Conrad, 1997). Alternatively, when respondents are asked a question about “health practitioners,” they may restrict their interpretation to physicians because physician is the prototypical health practitioner (Schaeffer &

Presser, 2003). Or, when asked to report “crimes,” respondents may restrict their interpretation to crimes committed by strangers because these are prototypical and leave out domestic crimes or crimes committed by relatives (Kindermann, Lynch, & Cantor, 1997; Lynch, 1996).

A potential problem with presenting examples is that rather than improving comprehension, they may limit recall only to the examples presented. This inhibiting effect is known as part-set cueing in the memory literature (see, for example, Roediger, 1974). In testimony before a House subcommittee, Scarr (1993) reported what appeared to be an inhibiting effect: adding “German” and dropping “English” as examples in the ancestry question in the 1990 Census resulted in a large increase in the number of people reporting German ancestry and a large decline in the number claiming English ancestry.

Subsequent research with examples, however, has not confirmed an inhibiting effect. For example, in a series of studies, Martin examined whether providing examples improved respondents’ comprehension of concepts (Martin, 2002; Martin, Sheppard, Bentley, & Bennett, 2007b). In both studies, the examples did not appear to restrict recall only to those groups mentioned, but instead broadened reporting to include more groups. For instance, in her 2002 paper, Martin compared two versions of an Hispanic origin question: one version of the question offered examples (Argentinian, Columbian, Dominican, Nicaraguan, Salvadoran, Spaniards); the second version did not provide any examples. Aside from Spaniards, the proportion of Hispanics writing in one of the example groups in response to the question with examples (7.8%) was not significantly different from the proportion of Hispanics

that wrote in one of the example groups in the question without examples (about 6%). However, the proportion reporting something other than one of the examples provided was significantly higher when examples were provided. Almost 9% of the Hispanics wrote in another Hispanic group when the question provided examples, despite the fact that these were not the same groups as listed in the examples. Only 4.2% of the Hispanics wrote in another Hispanic group in response to the question without examples.

Tourangeau, Conrad, Couper, Redline, & Ye (2009) experimented with providing examples in both the ancestry question and in a number of food frequency questions. Here again, there was no evidence that providing examples inhibited reporting. Respondents reported consuming an average of 6.9 servings of poultry and vegetables in a typical week when asked questions without examples versus consuming on average of from 6.9 to 8.2 servings of poultry and vegetables when asked questions with examples. In a follow-up analysis, Tourangeau, Conrad, Couper, & Ye (2010) found that respondents' answers to questions with examples depended on the type of examples given. The highest level of reporting, on average, occurred when the examples were relatively frequent, non-central instances of the category. So, for example when asked an open-ended question, respondents did not list French fries as a typical vegetable, but when French fries was among the examples mentioned, the average number of servings of vegetables that the respondents reported increased.

1.3 Obstacles to the Use of Clarification

Respondents either are not aware that they need clarification or not motivated enough to obtain it. Obtaining clarifying information may not be on the respondent's critical path², that is, respondents do not view the clarifying information as essential to their answering the questions, and they ignore it, sometimes knowingly, sometimes unknowingly (Conrad, Couper, Tourangeau, & Peytchev, 2006).

1.3.1 Presumption of Interpretability

One reason respondents may ignore clarifying information is that there may be a mismatch between respondents' everyday sense of a concept and the survey's use of it. When this occurs, respondents may rely on their everyday sense of a concept rather than the survey's technical definition (Tourangeau, Conrad, Arens, Fricker, Lee, & Smith, 2006). Clark & Schober (1992) suggest that this is one of the consequences of the presumption of interpretability, that is, respondents tacitly assume that surveyers chose wording that the respondent will quickly understand. As a result, respondents may fail to see when the surveyor is using everyday terms differently from the way respondents typically use them.

A compelling example of this comes from vignette research with the concepts of residency and disability, in which it was found that simply providing respondents with a definition for residency or disability did not improve respondents' classifi-

²Critical path analysis originated as a tool for project management to determine which activities were critical to completing a project, and refers to the shortest amount of time necessary to complete a project (e.g., Gray, John, & Atwood, 1993).

cations of the vignettes dramatically (Tourangeau, Conrad, Arens, Fricker, Lee, & Smith, 2006). When respondents were offered a definition that did not differ from their everyday understanding of the concept, the definitions were largely unnecessary. And when respondents were offered a definition that did differ from their everyday understanding of the concept (a technical definition), the definitions were largely ineffective – respondents answered in terms of their everyday understanding of the concept anyway. The researchers conclude that when a survey makes use of a technical sense of a concept, it may take collaboration with an interviewer to effectively convey such a definition. However, the technical definitions in this study were long and complex: each was composed of ten rules. Thus, it is not clear whether these findings generalize to all technical definitions or to only those that are long and complex.

In a related line of vignette research, Gerber, Wellens, & Keeley (1996) refer to rules that conformed to respondents' everyday sense of a concept as “intuitive rules” and showed that these rules had no effect on respondents' answers. For example, reminding respondents to include permanent household members who are temporarily away had no effect. Like Tourangeau and his colleagues (2006), Gerber and her colleagues found that respondents did not need to be given definitions in these cases. The researchers also studied “counterintuitive situations” (e.g., commuter workers who spend four days a week away in another state), and concluded that respondents were primarily influenced by their own definitions in classifying these vignettes. In other words, respondents' own sense of the concept seemed to win out over the survey's, although the researchers conclude that there did appear

to be a small benefit attached to presenting respondents with counterintuitive rules.

The studies by Tourangeau and his colleagues (2006) and Gerber and her colleagues (1996) were conducted in the laboratory with vignettes and were able to examine whether the vignettes were classified correctly or not. Conrad, Couper, Tourangeau, & Peytchev (2006), on the other hand, conducted a Web survey in which they did not have measures of accuracy. Instead, they collected paradata, which allowed them to determine whether respondents consulted definitions for survey terms or not. The researchers asked questions with technical (e.g., polyunsaturated fatty acid) or non-technical terms (e.g., vegetables) regarding food intake. Only a small percentage (17%) of the respondents consulted the definitions at all. Of those who consulted definitions, a minority (11%) requested definitions for the non-technical concepts. In comparison, definitions for technical concepts were requested a majority of the time (89%).

Altogether, this body of research shows the problem with using ordinary terms to describe technical concepts; however, what approach should be used instead is still not perfectly clear. Tourangeau, Conrad, Arens, Fricker, Lee, & Smith (2006) suggest two possibilities: one is to use a collaborative approach to convey the definitions and the other is to substitute a technical term for the ordinary term, e.g., use the term “enumeration unit” rather than “residence.” The fact that Conrad, Couper, Tourangeau, & Peytchev (2006) found that respondents were more likely to access definitions for technical rather than ordinary terms, and that their answers changed when they did so, provides some evidence in support of the second suggestion.

1.3.2 Conversational Order

Houtkoop-Steenstra (2002) makes the point that the effect of clarifying information — whether in the form of definitions, instructions, or examples — may depend on where the information is provided. Questions often begin by introducing the topic, then provide the clarifying information (definitions, instructions, or examples), and finish with the question itself.

Introduction, indicating something to talk about . . . :

Now we would like to talk about

. . . and what the topic is about:

your possible future plans with respect to such courses or education.

Clarifying Information:

They may be either fixed plans or vague ideas.

Question:

Do you have any plans for the COMING YEAR to take any course or education?

However, these components can and often are rearranged in surveys. In one common rearrangement, the clarifying information follows the question, as in the example below:

Introduction, indicating something to talk about:

Now we would like to ask you

Question:

do you have any plans for the COMING YEAR to take any course or education?

Clarifying Information:

They may be either fixed plans or vague ideas.

In conversational analysis, questions and answers are known as adjacency pairs (Schegloff & Sacks, 1973). Sachs, Schegloff, & Jefferson (1974) proposed a model for conversational turn-taking between such pairs. The researchers note that some conversational units seem to have points of completion that can be anticipated before they occur, such as the end of a question. Furthermore, these end points are identified through the use of the words and their arrangement, not their tone (De Ruiter, Mitterer, & Enfield, 2006). This may explain why respondents in interviewer-administered surveys interrupt the reading of the clarifying information when it follows the question (e.g., Houtkoop-Steenstra, 2002; Oksenberg, Cannell, & Kalton, 1991; Van der Zouwen & Dijkstra, 2002). Respondents seem to anticipate the end of the interviewer's turn (the reading of the question) and the beginning of their turn (to answer), and as a result, they stop processing the clarifying information. Similarly, it has been shown that respondents tend to answer questions without having read the clarifying information when the clarifying information follows the question in written forms (Jansen & Steehouder, 1992). Consequently, researchers,

such as Schaeffer & Presser (2003), advise against placing clarifying information after the question.

Respondents could be interrupting, however, because the clarifying information does not apply to them. If this is the case, the interruptions should not introduce error into the survey estimates. Without further direct evidence that these interruptions actually lead to biased estimates, we simply do not know whether it really matters whether the clarifying information comes before or after the question.

There is some evidence that placing clarifying information after the question in written forms may have an effect on survey estimates. Martin, Gerber, & Redline (2004) reported on three experiments conducted in Census 2000. In one of the experiments, the residency question was redesigned with the aim of improving within-household coverage. Many simultaneous changes were made to the experimental version of the question, but a change of particular interest was moving the definition from after the question-and-answer box to before it. This was done in an attempt to encourage respondents to read the definitions. Nonresponse to this question was significantly lower in the form in which the definition preceded the question-and-answer box. In addition, coverage improvements occurred for Hispanics, who are known to have relatively high rates of omission in the census. However, because numerous changes were made at once, it is difficult to say whether placing the definitions before the question was responsible for these positive results.

In a similar manipulation with skip instructions, Christian & Dillman (2004) found that placing a skip instruction “If you haven’t had many one-on-one meetings, skip to Question 9” before the response options resulted in 26% of the respondents

not responding, which presumably meant they skipped correctly. When the instruction followed the response options, this percentage dropped to 5%, which presumably meant they did not skip correctly.

1.3.3 Other Factors Affecting the Use of Clarification

Another reason why respondents may disregard clarifying information is that it is hard to access the clarifying information. For example, Conrad, Couper, Tourangeau, & Peytchev (2006) found that respondents rarely requested definitions by clicking a link (only 14% of the respondents requested such definitions). In line with this, Conrad, Couper, Tourangeau, & Peytchev (2006) found that the number of requests for definitions dropped as the number of clicks required increased (36% of the time respondents abandoned their request for a definition after the first click). Similarly, Tourangeau, Conrad, Arens, Fricker, Lee, & Smith (2006) found that only about a fifth of the respondents clicked to access vignettes in their study.

Respondents may also disregard clarifying information because it does not attract their attention visually. Redline and her colleagues demonstrated that visual cues could be used to heighten the performance of branching instructions (Redline & Dillman, 2002; Redline, Dillman, Dajani, & Scaggs, 2003). Still, from 14 to 21% of the time, depending on the design of the branching instruction, the visual manipulations did not work.

The implication of this research is that we do not sufficiently understand how to attract respondents' attention through the use of visual cues. For example, does

placing clarifying information in a different font from that of the question, such as italics, attract respondents' attention or signal that the clarifying information is optional? According to one theory, incongruity in font catches a reader's attention and motivates processing (McCarthy & Mothersbaugh, 2002). This theory is based on the presumption that information operates under the Gricean maxims (Grice, 1975). Since different fonts are not necessary for understanding the text, the implicature drawn by readers may be that the different fonts were chosen to convey or highlight the importance of the text.

According to a second theory, however, changing font violates one of the Gestalt grouping laws that respondents are thought to follow in interpreting visual information in questionnaires, the grouping law of similarity. This principle states that similar objects are more likely to be perceived as a cohesive unit (Wertheimer, 1938). Jenkins & Dillman (1997) discuss this principle as it relates to surveys. Under this competing theory, respondents are less likely to attend to the clarifying information when it is in a different font as the question because it does not appear to be part of the question. However, it is also possible that it does not matter whether the clarifying information is italicized or not.

Although clarifying information can reduce ambiguity and vagueness, it can also introduce error if it leads to an increase in complexity and working memory overload (e.g., Fowler, 1992, 1995, p. 17; Tourangeau, Rips, & Rasinski, 2000, p. 45). An example of this comes from Fowler (1992), in which a question that respondents appeared to have difficulty comprehending was subsequently clarified:

Original: During the past twelve months, that is, since January 1, 1987, about how many days did illness or injury keep you in bed more than half of the day? (Include days while an overnight patient in a hospital.)

Clarified: The next question is about extra time you have spent in bed because of illness or injury (including time spent in the hospital). During the past twelve months, since July 1, 1987, on about how many days did you spend several extra hours in bed because you were sick, injured, or just not feeling well?

The clarified version resulted in 30% of the interviews exhibiting inadequate answers compared to 7% for the original. The “clarification” also did not reduce the proportion of respondents asking for clarification. Complexity appears to be the result of two properties: length and complicated syntax (e.g., Bishop & Smith, 2001; Holbrook, Krosnick, Moore, & Tourangeau, 2007; Holbrook, Cho, & Johnson, 2006; Graesser, Bommarreddy, Swamer, & Golding, 1996; Graesser, Kennedy, Wiemer-Hastings, & Ottati, 1999; Graesser, Cai, Louwerse, & Daniel, 2006; Yan & Tourangeau, 2008). Thus, researchers have long suggested that complexity may be reduced by asking a series of shorter, simpler questions instead of a single complicated question (Conrad & Couper, 2004; Conrad & Schober, 2000; Couper, 2008, p. 289; Fowler, 1995, pp. 13-20; Jenkins & Dillman, 1997; Schaeffer & Presser, 2003; Sudman, Bradburn, & Schwarz, 1996, p. 31; Suessbrick, Schober, & Conrad, 2000; Tourangeau, Rips, & Rasinski, 2000, pp. 38-40, 61). Considering the number of

times this strategy has been suggested, it is surprising that there is little empirical research supporting it.

1.4 Sensory Channel Effects

Several studies have demonstrated that answers to survey questions differ across modes (for a review, see Tourangeau, Rips, & Rasinski, 2000). Numerous models have been advanced for explaining these mode effects (e.g., DeLeeuw & Van der Zouwen, 1988; Tourangeau, Rips, & Rasinski, 2000; Dillman, 2000). For present purposes, a key issue is whether sensory channels (aural versus visual) make different demands on comprehension. Although this has certainly been postulated, the answer to this question is largely unknown because so many of the relevant studies compare self-administered surveys (mail or Web) with interviewer-administered surveys (telephone or in-person) (e.g., Bishop, Hippler, Schwarz, & Strack, 1988; Chang & Krosnick, 2009; Christian, Dillman, & Smyth, 2007; Fricker, Galesic, Tourangeau, & Yan, 2005; Smyth, Christian, & Dillman, 2008). Sensory channel is typically confounded with the presence of an interviewer in these studies.

Several papers (Clark & Brennan, 1991; Just & Carpenter, 1980; Osada, 2004; Rayner & Clifton, 2009; Schwarz, Strack, Hippler, & Bishop, 1991a) suggest that written language differs from spoken language in a number of important ways:

1. reading is visual and spatial, whereas speech is auditory and temporal;
2. readers can control the pace of input, whereas listeners usually can

not;

3. readers can preview written input (i.e., see that something is present or lies ahead without necessarily reading it), whereas listeners can not;
4. readers can review written input, whereas listeners must rely much more heavily on working memory;
5. text is supplemented by visual cues, such as color, shape, and location, whereas speech is supplemented by aural cues, such as stressed words and variations in pace; and
6. readers can repair errors privately, whereas speakers cannot.

How do these properties relate to respondents' understanding of survey questions?

1.4.1 Channel and the Presumption of Interpretability

Respondents tacitly assume that the researchers have chosen wording that they can understand (Clark & Schober, 1992). This may explain why respondents rely on their own sense of a word rather than apply the definition provided (e.g., Tourangeau, Conrad, Arens, Fricker, Lee, & Smith, 2006). This tendency to ignore definitions and rely on one's prior sense of a concept may be greater in aural surveys where respondents cannot control the rate at which information is received and there is greater time pressure to respond quickly (Schwarz, Strack, Hippler, & Bishop, 1991a). This suggests that comprehension errors may be greater in the aural channel

than the visual.

1.4.2 Channel and Order

Because words rather than tone signal the end of a speaker's turn (De Ruiter, Mitterer, & Enfield, 2006), respondents may be as likely to stop the reading of clarifying information when it follows a question in the visual channel as they are to interrupt an interviewer in the aural channel. However, the placement of the clarifying information may matter more in aural surveys because listeners may not be able to take in information at a pace that matches their internal comprehension (Just & Carpenter, 1980). Also, listeners can only review information in the aural channel by recalling it or asking an interviewer to repeat it. Asking interviewers to repeat information appears to be rare (Schwarz, Strack, Hippler, & Bishop, 1991a) probably because, as Clark & Brennan (1991) point out, speakers cannot repair errors privately. Respondents may be too embarrassed to admit they were not able to keep pace with a speaker, which may discourage them from asking questions to be repeated (Conrad, Couper, Tourangeau, & Peytchev, 2006).

By contrast, readers can repair their misunderstandings privately; they can go back and reread the information they missed (Just & Carpenter, 1980). Thus, researchers have speculated that it may matter less where clarifying information is placed in a visual survey than an aural one (Martin, Hunter Childs, DeMaio, Hill, Reiser, Gerber, Styles, & Dillman, 2007a). Many surveys administered through the visual channel, such as the American Community Survey, place clarifying in-

formation after the questions rather than before. However, unanswered questions surround this issue. Do readers stop reading clarifying information when it is placed after the question in the visual channel (similar to listeners interrupting the reading of the question in the aural channel)? If so, do these same readers repair consequent misunderstandings by returning to re-read it?

1.4.3 Channel and Other Factors

Changing the visual appearance of the clarifying information might lead to differences between the sensory channels by differentially affecting the processing of the information. For example, if the number of respondents who read the clarifying information in the visual channel is reduced because the clarifying information is changed from a bold to italic print, this would result in response differences between the channels.

In addition, the aural channel may promote working memory overload relative to the visual channel so that complicated questions might show larger cross-channel differences than simpler questions. There seems to be general agreement that listening is harder and reading is easier when the information to be taken in is complex (Clark & Brennan, 1991; Osada, 2004; Tourangeau, Rips, & Rasinski, 2000, p. 302), and it has long been asserted that questions in a telephone interview should be “simpler” than those administered face-to-face (Groves, 1989, p. 520). Fricker, Galesic, Tourangeau, & Yan (2005) reported some empirical support for this interaction between channel and question complexity. They found a significant mode by item

type interaction, with differences between a telephone and Web survey smallest for the least demanding question forms (true/false forms) and largest for the most demanding ones (open-ended questions). Multiple choice questions fell between these two (true/false and open-ended). Similarly, Bishop & Smith (2001) found that response order effects were more likely to occur in complex questions than simple ones administered aurally.

A study from the communication literature illustrates why complicated questions may demonstrate larger cross-channel differences than simpler ones. In this particular study, tutors instructed students on assembling a pump over the telephone versus through the use of keyboard conversations (Cohen, 1984). Clark & Brennan (1991) make the point that there were many more separate exchanges over the telephone than in keyboard conversations. Clark & Brennan (1991) referred to these as installments, and proposed that dividing a presentation into installments is based on the tacit recognition that people have limited immediate memory spans.

A summary of the major linguistic differences between written and spoken language encapsulates the previous discussion: “In spoken language, idea units tend to be shorter, with simpler syntax, whereas written units tend to be more dense, often using complex syntax, such as dependent and subordinate clauses, to convey more information” (Buck, 2001; see also Osada, 2004). Although readers may potentially be better equipped to understand denser language than listeners, the question is whether they would not benefit from simpler language as well?

1.5 Summary of the Literature

Belson (1981) found that respondents have difficulty comprehending questions in surveys. As with any single piece of research, Belson's study had methodological limitations; however, a body of research, both theoretical and empirical, has accumulated since Belson's seminal study suggesting that respondents often do experience comprehension problems (e.g., Conrad & Schober, 2000; Tourangeau, Conrad, Arens, Fricker, Lee, & Smith, 2006).

Ordinary concepts may be polysemous or exhibit other forms of ambiguity (Klein & Murphy, 2001; Murphy, 1997; Nunberg, 1979). They may also be vague or exhibit conceptual variability (Pinkal, 1995). Further, concepts that seem straightforward may not map onto the respondent's situation in a straightforward way (e.g., Conrad, Schober, & Coiner, 2007). It seems that the broader the concept, or the more that it involves gradations, the more likely it is to be vague (Ruhl, 1989). When this is the case, respondents need clarification regarding which instances should be included and which ones should be excluded. Otherwise, they may expand or restrict the category of interest in ways the researchers did not intend.

Being broad, however, does not guarantee comprehension problems. Centrality plays a critical role here (Tourangeau, Conrad, Arens, Fricker, Lee, & Smith, 2006). Central instances are ones that respondents easily recognize as fitting their definitions of a category, and consequently, their membership in the category appears to be less ambiguous than that of the more peripheral or non-central instances. For example, tables and chairs are central instances of the concept of furniture. Non-central

instances fit the concept less well and therefore, their membership may be more ambiguous. For example, patio furniture is a non-central instance of furniture, and its membership in that category is likely to be more ambiguous. Respondents appear to need less clarification regarding central instances than non-central instances.

However, simply providing respondents with clarification will not help if respondents do not use it— that is, they do not view it as essential to their answering the survey questions correctly. One method that has clearly increased respondents' use of clarifying information is the “conversational method of interviewing,” in which interviewers are allowed to intervene with respondents to clarify understanding (e.g., Conrad & Schober, 2000; Schober & Conrad, 1997). Interviewers are equipped with knowledge regarding what the survey concepts mean and are encouraged to impart this knowledge to respondents as they deem necessary. This method has been shown to reduce comprehension errors when it is compared to the standardized method of interviewing in which interviewers are not allowed to clarify concepts for respondents.

A number of things seem to prevent respondents from seeing clarifying information as essential to answering the questions. Respondents may not think they need clarification, as when the surveys uses an everyday term (like “residence”) in a technical way (e.g., Tourangeau, Conrad, Arens, Fricker, Lee, & Smith, 2006). Under such circumstances, respondents appear to rely on their own sense of a word or concept and ignore the survey's definition.

Putting the clarifying information after the question may reduce its effectiveness. Conversational analysis finds problems when the clarifying information is

placed out of order: respondents may interrupt a question and answer before the interviewer is able to read the clarifying information (e.g., Houtkoop-Steenstra, 2002; Oksenberg, Cannell, & Kalton, 1991; Van der Zouwen & Dijkstra, 2002). However, it has not yet been shown that such interruptions actually lead to changes in respondents' answers to the questions. I think it does, however, and thus I hypothesize that respondents are less likely to process clarifying information placed after a question than before.

A third barrier to the use of clarifying information is the inaccessibility of the information to respondents. For example, clarifying information that requires respondents to click on a link is less accessible to respondents than information that is visible without any mouse clicks (Conrad, Couper, Tourangeau, & Peytchev, 2006). An unanswered research question surrounding this issue is if making the clarifying information more accessible increases its use, and further whether there are any cues that suggest to respondents that the clarifying information is essential. For example, some surveys place the clarifying information in italics. Does this help, hinder, or have no effect on its being read? I hypothesize that respondents are more likely to view clarifying information placed in italics as optional, and therefore, less likely to process it than when it is placed in the same font as the question.

Many researchers believe that providing clarification improves comprehension, but that such clarification is less effective when it is long and complex (e.g., Fowler, 1992, 1995, p. 17; Tourangeau, Rips, & Rasinski, 2000, p. 45). When the clarification is elaborate, many researchers recommend incorporating the clarification into the questions and asking multiple questions instead, but little empirical research exists

to support this recommendation (Conrad & Couper, 2004; Conrad & Schober, 2000; Couper, 2008, p. 289; Fowler, 1995, pp. 13-20; Jenkins & Dillman, 1997; Schaeffer & Presser, 2003; Sudman, Bradburn, & Schwarz, 1996, p. 31; Suessbrick, Schober, & Conrad, 2000; Tourangeau, Rips, & Rasinski, 2000, pp. 38-40,61). I agree with these researchers and thus I hypothesize that respondents are less likely to process clarifying information when it is dense (one question with elaborate clarification) than when it is less dense (multiple simple questions that incorporates the clarification).

Finally, there are theoretical reasons to think that the sensory channel in which the clarification is administered will influence the degree to which the clarification is understood. The visual channel differs from the aural channel in ways that can improve respondents' ability to comprehend clarifying information (Clark & Brennan, 1991; Just & Carpenter, 1980; Osada, 2004; Rayner & Clifton, 2009; Schwarz, Strack, Hippler, & Bishop, 1991a). Thus, I hypothesize that readers will comprehend clarifying information better than listeners. Furthermore, the degree to which the sensory channel influences respondents' answers is also expected to interact with how the clarification is presented.

Chapter 2

Clarifying Instructions in a Web Survey

2.1 Introduction

Respondents seem to have difficulty understanding survey questions (Belson, 1981; Schober & Conrad, 1997). There are multiple reasons why respondents may misunderstand questions. For example, Tourangeau, Rips, & Rasinski (2000, pp. 34-61) distinguish seven types of comprehension errors. Two of them are the focus of this chapter: lexical ambiguity and vague concepts. Lexical ambiguity occurs when a word has more than one meaning. One form of lexical ambiguity is polysemy, or words that have different, but related meanings (e.g., Klein & Murphy, 2001; Lakoff, 1987; Murphy, 1997; Nunberg, 1979). For example, teachers have been known to report that they have from twenty to thirty “children,” illustrating that the word “children” can mean “young people” or “students” rather than “offspring” (Sudman, Bradburn, & Schwarz, 1996, pp. 60-61).

Vague concepts are another important problem in surveys; these are concepts that have unclear boundaries or that permit multiple variations on a single meaning (Pinkal, 1995; Ruhl, 1989). In a study that examined a Current Population Survey supplement, Suessbrick, Schober, & Conrad (2000) showed that “smoking cigarettes” meant “only cigarettes you finished” to some respondents, but “even just one puff” to others. These concepts all have the same basic meaning, but they

vary enough to produce different answers.

A third reason respondents may misunderstand questions stems from the fact that concepts that seem straightforward may not map onto a respondent's situation in a straightforward way. Numerous studies have demonstrated that respondents whose situations clearly fit the questions have less difficulty answering questions correctly (Conrad & Schober, 2000; Conrad, Schober, & Coiner, 2007; Gerber, Wellens, & Keeley, 1996; Schober & Conrad, 1997; Schober & Bloom, 2004; Tourangeau, Conrad, Arens, Fricker, Lee, & Smith, 2006). For example, a family in which everyone lives at home year round will have less trouble answering a question about the number of people living in the household than a family who has a son or daughter living in a dormitory during the school year, but who is home the rest of the year. Tourangeau, Conrad, Arens, Fricker, Lee, & Smith (2006) suggest that such a mapping reflects an imperfect fit between the survey's definition of a concept and the respondent's particular circumstances.

Methods for reducing comprehension problems in survey questions include providing clarification in the form of definitions, instructions, or examples. For instance, allowing interviewers the freedom to give definitions has been shown to improve the accuracy of respondents' answers (Schober & Conrad, 1997, see also Conrad & Schober, 2000). Definitions are most effective when they are presented along with the questions without requiring any action from the respondent, such as more clicks (Conrad, Couper, Tourangeau, & Peytchev, 2006; Conrad, Schober, & Coiner, 2007); instructions are more likely to be followed when they are made more visible (e.g., Redline, Dillman, Dajani, & Scaggs, 2003). Finally, providing examples

(subcategories) that are meant to illustrate the concept being asked about has been shown to alter understanding (Martin, 2002; Tourangeau, Conrad, Couper, Redline, & Ye, 2009; Tourangeau, Conrad, Couper, & Ye, 2010).

This chapter investigates several factors that may affect respondents' use of material intended to clarify a question. For one, conversational analysis has shown that respondents anticipate the end of a question and are more likely to interrupt when clarifying information is placed after a question than when it comes before a question (e.g., Houtkoop-Steenstra, 2002; Oksenberg, Cannell, & Kalton, 1991; Van der Zouwen & Dijkstra, 2002). The position of the information may matter more in the aural channel. However, it has never been firmly established that these interruptions actually had an effect on respondents' answers or that a similar phenomenon occurs when questions are presented visually. It could be that respondents interrupt because their situations are straightforward (Conrad & Schober, 2000; Conrad, Schober, & Coiner, 2007; Schober & Conrad, 1997; Schober & Bloom, 2004) and the clarifying instructions do not apply to them. And, it may be that placing the clarifying instructions after a question is genuinely more problematic in the aural channel than the visual (Martin, Hunter Childs, DeMaio, Hill, Reiser, Gerber, Styles, & Dillman, 2007a); respondents can see that additional text follows the question in the visual channel and they can read it and reread it if need be (Just & Carpenter, 1980). Thus, the research questions are whether placing the clarifying instructions before the question improves comprehension and whether this improvement is larger in the aural channel than in the visual channel?

Other characteristics of the clarifying information may affect whether respon-

dents process the information in the visual channel. According to one theory, changing font violates the Gestalt law of similarity (Wertheimer, 1938), which states that perceptually similar objects are more likely to be perceived as a cohesive unit. Jenkins & Dillman (1997) describe how this principle relates to surveys. Under this theory, when the clarifying information is in the same font as the question, the respondent will be more likely to read it because it will be seen as part of the question (i.e., on the critical path). According to a second theory, however, changes in font catch a reader's attention and motivate processing (McCarthy & Mothersbaugh, 2002). If this is true, italicized instructions may increase the chances that the clarifying information will be read. Thus, the research question is: does placing the clarifying information in the same font as the question or italicizing it increase its chances of being processed?

Finally, although it has been demonstrated that providing clarification can improve comprehension, it is possible that when the clarifying instructions are long and complex, they can tax working memory. As a result, it may be better to recommend incorporating the clarifying instructions into the questions, and possibly asking multiple questions rather than a single question with clarifying instructions (Conrad & Couper, 2004; Conrad & Schober, 2000; Couper, 2008, p. 289; Fowler, 1995, pp. 13-20; Jenkins & Dillman, 1997; Schaeffer & Presser, 2003; Sudman, Bradburn, & Schwarz, 1996, p. 31; Suessbrick, Schober, & Conrad, 2000; Tourangeau, Rips, & Rasinski, 2000, pp. 38-40, 61). However, surprisingly little empirical research has tested this recommendation.

This chapter describes an experiment designed to investigate where and how

to present clarifying information so that respondents recognize it as essential to their answering survey questions correctly. The experiment examines whether the addition of clarification improves comprehension, whether placing the clarification before the question is better than putting it afterwards, whether putting the clarification in the same font as the question helps, or whether incorporating the clarifications into a series of questions has greater effect than providing instructions on how to answer.

2.2 Methods

2.2.1 Study Items

In this experiment, eight study questions, based on items from major federal surveys, were administered as part of a Web survey. The respondents to the Web survey had been recruited from an area probability sample. Half of the questions in this experiment were patterned after questions in the American Community Survey (ACS). These questions asked about the number of people living in the sample household (from the cover page of the mail 2008 ACS form), the number of rooms and the number of bedrooms in the sample dwelling (Items 7a and 7b from the housing section of the 2008 form), and whether the respondent worked last week (Item 28 from the person section of the form). At the time these questions were selected, they were in somewhat different forms in the different ACS modes (paper versus face-to-face), making them good candidates for this study. I converted the “did you work last week” question (Item 28) into a question asking for the number

of hours worked last week. The remaining questions were based on items from the Current Population Survey (CPS), the Consumer Expenditure Survey (CES), and the American Time Use Survey (ATUS). The target items, thus, included these eight questions:

1. How many people are currently living or staying at this address?
2. How many pairs of shoes do you own?
3. How many coats and jackets do you own?
4. Last week, how many hours, if any, did you work for either pay or profit?
5. In the past year, how many times, if any, were you away from home on a trip?
6. In the past year, how many furniture purchases, if any, did you make?
7. How many bedrooms are in this house, apartment, or mobile home?
8. How many other separate rooms are in this house, apartment, or mobile home?

All of these items ask about concepts that may require further clarifying instructions to understand correctly, especially if the questions map onto the respondent's situation in a complicated way. For the purposes of this study, the clarifying instructions for each of the items instructed respondents to exclude some of the subclasses that they might otherwise be likely to include in the category of interest. For example, the clarifying instructions in the *shoe* question instructed respondents to exclude boots, sneakers, athletic shoes, and bedroom slippers. The instructions were written so that, if they were followed, respondents would, on average, give

lower answers. This allowed me to predict in advance the direction in which the answers should move. To increase the magnitude of the effects, I tried to choose clarifying instructions that would apply to many respondents. For example, I asked respondents to exclude sneakers in the *shoe* question under the assumption that many people owned sneakers (see Appendix A.1. for the complete questionnaire, including all clarifications).

The questions underwent numerous iterations, a couple of expert reviews, and a small pretest (with 12 respondents chosen for convenience).

2.2.2 Experimental Conditions

This experiment compared the effectiveness of three main conditions – no clarifying instructions, clarifying instructions presented in four different ways, and a multiple question approach. Embedded within the four different ways of presenting clarifying instructions were two orders of presenting the instructions (after/before) crossed with two font styles (same font as the question/italics), for an overall design with six groups. The six methods are detailed below.

Method 1: Question with No Clarifying Instructions. As shown in Table 2.1, this method served as the base comparison. This method consisted of a single question with no clarifying instructions. The question was presented in bold and a preface in the same bold font preceded the question.

Method 2: Question with Clarifying Instructions After. This method consisted of a question with clarifying instructions presented after the question. The preface,

the question, and the clarifying instructions were all presented in the same bold font.

Method 3: Question with Clarifying Instructions Before. This method first presented the preface, then the clarifying instructions, and finally the question itself. All information was placed in bold.

Method 4: Question with Italicized Clarifying Instructions After. This was the same as method 2, except that the clarifying instructions were in italics rather than bold.

Method 5: Question with Italicized Clarifying Instructions Before. This was the same as method 3, except that the clarifying instructions were in italics rather than bold.

Method 6: Multiple Questions with Clarifying Instructions Incorporated. This method was based upon an approach recommended by Fowler (1995). In this method, one question was separated into a series of questions, each of which incorporated one of the clarifying instructions.

2.2.3 Accessing the Effectiveness of the Clarifying Instructions

Ideally, I would have compared respondents' answers to administrative records or some other "gold standard." However, given the impracticality of this, I assessed the effectiveness of the instructions in two ways. First, as the study items were written, numeric responses should have been greater if the respondents ignored the

Table 2.1: Examples of the No Clarifying Instruction and Alternative Clarifying Instruction Methods

<p>Method 1</p> <p>2A. The next question is about your footwear.</p> <p>How many pairs of shoes do you own?</p> <p>Number of pairs of shoes</p> <input data-bbox="410 716 532 762" type="text"/>	<p>Method 2</p> <p>2B. The next question is about your footwear.</p> <p>How many pairs of shoes do you own? For the purposes of this question, do not include boots, sneakers, athletic shoes, or bedroom slippers. Include sandals, other casual shoes, and dress shoes. If you do not own a pair of shoes (as we have defined them), enter “0.”</p> <p>Number of pairs of shoes</p> <input data-bbox="737 888 859 934" type="text"/>	<p>Method 3</p> <p>2C. The next question is about your footwear.</p> <p>For the purposes of this question, do not include boots, sneakers, athletic shoes, or bedroom slippers. Include sandals, other casual shoes, and dress shoes. If you do not own a pair of shoes (as we have defined them), enter “0.” How many pairs of shoes do you own?</p> <p>Number of pairs of shoes</p> <input data-bbox="1071 863 1193 909" type="text"/>
<p>Method 4</p> <p>2D. The next question is about your footwear.</p> <p>How many pairs of shoes do you own? For the purposes of this question, do not include boots, sneakers, athletic shoes, or bedroom slippers. Include sandals, other casual shoes, and dress shoes. If you do not own a pair of shoes (as we have defined them), enter “0.”</p> <p>Number of pairs of shoes</p> <input data-bbox="410 1356 532 1402" type="text"/>	<p>Method 5</p> <p>2E. The next question is about your footwear.</p> <p><i>For the purposes of this question, do not include boots, sneakers, athletic shoes, or bedroom slippers. Include sandals, other casual shoes, and dress shoes. If you do not own a pair of shoes (as we have defined them), enter “0.”</i> How many pairs of shoes do you own?</p> <p>Number of pairs of shoes</p> <input data-bbox="737 1331 859 1377" type="text"/>	<p>Method 6</p> <p>2F1. The next question is about your footwear.</p> <p>How many pairs of shoes do you own?</p> <p>Number of pairs of shoes</p> <input data-bbox="1102 1178 1224 1224" type="text"/> <p>[If 2F1 > 0:]</p> <p>2F2. When you reported the pairs of shoes that you own, how many pairs of boots, sneakers, athletic shoes, or bedroom slippers, if any were included?</p> <p>Number of pairs of shoes</p> <input data-bbox="1096 1440 1218 1486" type="text"/> <p>[If 2F1=2F2, do not ask 2F3.]</p> <p>2F3. When you reported the pairs of shoes that you own, how many pairs of sandals, other casual shoes, or dress shoes, if any, were included?</p> <p>Number of pairs of shoes</p> <input data-bbox="1096 1696 1218 1743" type="text"/>

clarifying instructions than if they attended to them. Second, I included a couple of follow-up questions to the *shoe* and the *hours worked* questions to evaluate whether respondents actually followed the clarifying instructions. These responses provide a second measure for determining which of the methods elicited the most accurate reporting (see Conrad & Schober 2000 for a similar method).

2.2.4 Data Collection and Sample

This experiment compares answers obtained under the six experimental versions of the items. Each respondent was randomly assigned to one version and received all of the study questions in that format. This experiment was administered from August 1 through October 31, 2009 as part of a Web survey administered to a probability sample. Details of the sample design and weighting procedures can be found in Tourangeau & Sakshaug (2010). The sample was assembled by Abt SRBI under a NSF Major Research Instrumentation grant to Jon Krosnick. A representative sample of addresses in the U.S. was selected from the U.S. Postal Service mailing list. Interviewers visited the selected households, randomly selected an adult member, conducted a brief face-to-face interview, and offered a free laptop and high-speed connection. In exchange, respondents agreed to complete a 30-minute Internet survey once a month for 12 months. Of the 1000 respondents recruited into the survey panel, 913 responded by the end of October 2009. The overall unweighted response rate for the initial 1000 recruits (AAPOR RR4) was 42.5% (Sakshaug, Tourangeau, Krosnick, Ackermann, Malka, DeBell, & Turakhia,

2009). Thus, the response rate for the August panel was 38.8% (91.3 x 42.5%).

The eight study items and two follow-up questions that made up this experiment were part of a larger questionnaire that included four other experiments in the month of August, and asked twenty seven questions in total. The experimental questions were dispersed throughout the questionnaire, as questions 1, 8, 9, 10, 12, 13, 17, 18, 26 and 27.

To maintain as much equivalence as possible between the multiple and single-question versions of the items, the multiple questions were presented on one screen. The multiple questions were presented dynamically, which means that only the relevant follow-up questions were administered based on the prior responses. For example, if a person reported having five persons in a household, all of whom were adults, they were not asked a follow-up question concerning how many children there were in the household.

2.3 Results

The analysis examines three outcome variables. The main outcome variable was the overall level of reporting in the answers to the study items. One comparison looked at the overall level of reporting in the three main conditions (no clarifying instructions, clarifying instructions and multiple questions) and a second focused on the level of reporting in the different versions of the clarifying instructions. The second outcome variable was based on respondents' answers to the two follow-up questions. This analysis assessed whether responses to the *shoes* and *hours worked*

questions were consistent with the clarifying instructions (whether these instructions were present or not). The third outcome variable was the amount of time it took respondents to answer each of the study questions in the clarifying methods only.

2.3.1 Responses to the Study Items

Respondents' answers to the study items were numeric, and some respondents reported extreme values. Values that were above the upper one percentile for each individual item were removed.¹

In the no clarifying instructions condition and the four conditions with clarifying instructions, the responses to each item were given in an answer to one question. In the multiple questions method, responses to each item were calculated from respondents' answers to the set of questions. For example, as shown in Method 6 of Table 2.1, answers to the *shoe* question were derived by subtracting the number of boots, sneakers, athletic shoes, and bedroom shoes in 2F2 from the number of shoes owned in 2F1. Negative values that resulted from this calculation were set to missing.

Figure 2.1 and Table 2.2 display the mean response for each of the items for the three main experimental conditions (no clarifying instructions versus clarifying instructions versus multiple questions). As can be seen in Figure 2.1, seven of

¹The cutoff and the number of values removed for each item was: residents > 8 (9 values); shoes > 100 (5 values); coats > 35 (9 values); hours worked > 70 (9 values); trips > 24 (9 values); furniture > 8 (9 Values); bedrooms > 6 (5 values); rooms > 12 (9 values). Removing these outliers did not change the overall outcome of the one-way ANOVAS reported below.

the eight items display the expected downward trend across these three groups. Only one item, *hours worked*, revealed a non-significant reversal between the no clarifying instruction and the clarifying instruction methods.² Of the seven items that displayed a downward trend, six were significantly different in an one-way ANOVA. One item, *furniture purchases*, moved in the direction of the expected downward direction; however, these differences were not significant.³

Table 2.2 provides the mean responses to each of the items for the three main conditions. To illustrate the overall pattern, I describe the results for the number of residents because this item exhibited close to the average percent reduction. Providing clarification in this item reduced the mean response by 20% compared to the no clarifying instructions condition (a mean of 3.0 versus a mean of 2.4). Asking multiple questions reduced the mean by 33% compared to the no clarifying instructions group (3.0 versus 2.0). These results suggest that providing clarification was less effective than asking multiple questions.⁴

²The other two pairwise comparisons for this item (no clarifying instructions versus multiple questions and clarifying instructions versus multiple questions) significantly decreased, as expected.

³Further examination of the furniture question revealed that a large proportion of the responses to the three main versions of the item (65 % in the no clarifying instruction version, 67 % in the one question with clarifying instruction versions and 71% in the multiple questions version) were zeros. Thus, a likely reason that the clarifying instructions did not have a larger effect was because so few people had bought any furniture in the first place.

⁴I also compared the mean response between each of the questions in the no clarifying instruction condition with those of the first question in the multiple question series using ANOVA. None were significantly different. Also, the standardized mean across all eight questions for the no clarifying instruction condition (0.4) was not any different from the standardized mean for all eight of the first questions in the multiple question series (0.8)($F(1, 329) = 1.52$, n.s.). This suggests that

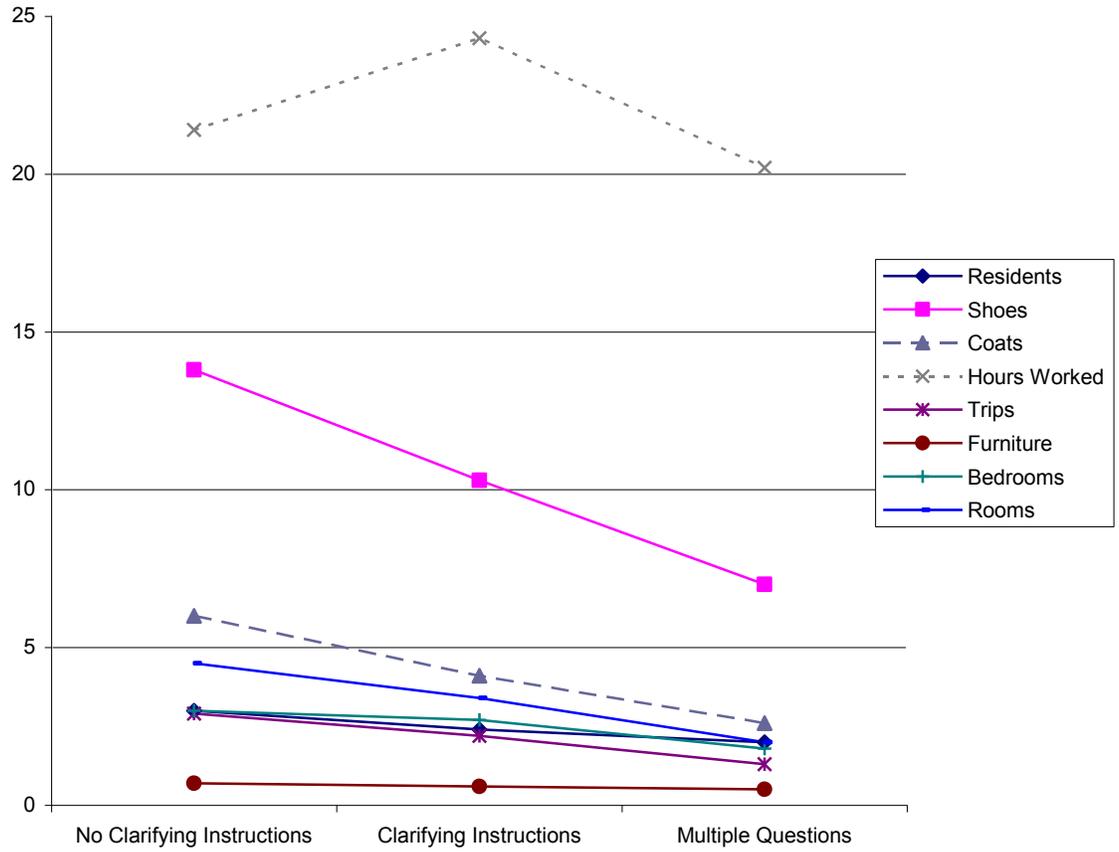


Figure 2.1: Mean Response to Items as a Function of the Main Conditions

Table 2.2: Mean Response (and Sample Sizes) by Main Condition and Item

Item	No Clarifying Instructions	Clarifying Instructions	Multiple Questions	Main Conditions
	Mean (n)	Mean (n)	Mean (n)	<i>F</i> -statistic
Residents	3.0 (174)	2.4 (572)	2.0 (148)	26.33***
Shoes	13.8 (176)	10.3 (576)	7.0 (151)	11.36***
Coats	6.0 (175)	4.1 (572)	2.6 (140)	30.43***
Hours Worked	21.4 (177)	24.3 (572)	20.2 (147)	3.02*
Trips	2.9 (174)	2.2 (575)	1.3 (149)	13.47***
Furniture	0.7 (177)	0.6 (574)	0.5 (152)	1.16 n.s.
Bedrooms	3.0 (177)	2.7 (573)	1.8 (153)	44.73***
Rooms	4.5 (174)	3.4 (569)	2.0 (145)	61.49***

Note: Values that were greater than the upper one percentile for each individual item were removed. *F* statistics are from one-way ANOVAS; all are based on 2 numerator degrees of freedom. *** $p < .001$; ** $p < .01$; * $p < .05$; n.s. denotes not significant

Narrowing our focus to the methods with clarifying instructions only, Table 2.3 displays the individual means for each of the eight study items as a function of the order of the information. The font variable had no effect on reports ($F(1, 577) = 0.0$, n.s.) and I focus on the order variable here. As can be seen, the overall pattern is in the expected direction, with lower means when the clarifying instructions come before the question than afterward. Five of the eight items clearly moved in this direction. One item, *shoes*, displayed a nonsignificant reversal, and two items (*trips* and *furniture*) were flat.

I again use the number of residents to illustrate the results. Placing the clarifying instructions before the question reduced the mean response by 8% compared to placing them after the question (from a mean of 2.5 residents to a mean of 2.3).

Because the individual effects were smaller here than they had been when comparing the three main conditions, the overall trend across all eight test items was less clear. I combined responses to all eight questions after I standardized the responses to each question to have a mean of 0 and a standard deviation of 1. Figure 2.2 displays the standardized means for each item.

The standardized means tend in the predicted direction. The dotted line represents the combined responses to all eight questions. The combined response respondents did not give zero answers to the initial question to avoid the followup questions. In addition, a chi-squared analysis revealed that there was no difference in the percentage of zeros given in response to all eight of the questions with no clarifying instructions (16.9 %) versus the percentage of zeros given in response to all eight of the first questions in the multiple question series (17.3 %) ($\chi^2(1, N = 330) = 0.03$, n.s.)

Table 2.3: Mean Response (and Sample Sizes) by Order of Presenting Clarifying Instructions and Item

Item	Order				F - statistic	
	Clarifying Instructions After		Clarifying Instructions Before			After vs. Before
	Mean	(n)	Mean	(n)		
Residents	2.5	(293)	2.3	(279)	3.54 †	
Shoes	10.2	(297)	10.4	(279)	0.04 n.s.	
Coats	4.3	(294)	3.8	(278)	1.80 n.s.	
Hours Worked	24.5	(292)	24.0	(280)	0.70 n.s.	
Trips	2.2	(295)	2.2	(280)	0.00 n.s.	
Furniture	0.6	(296)	0.6	(278)	0.00 n.s.	
Bedrooms	2.8	(294)	2.5	(279)	9.70 **	
Rooms	3.6	(292)	3.2	(277)	4.80 *	

Note: Values that were greater than the upper one percentile for each individual item were removed. F -statistics for the individual items are from 2-way ANOVAS; all are based on 1 degree of freedom.

** $p < .01$; * $p < .05$; † $p < .1$; and n.s. denotes not significant

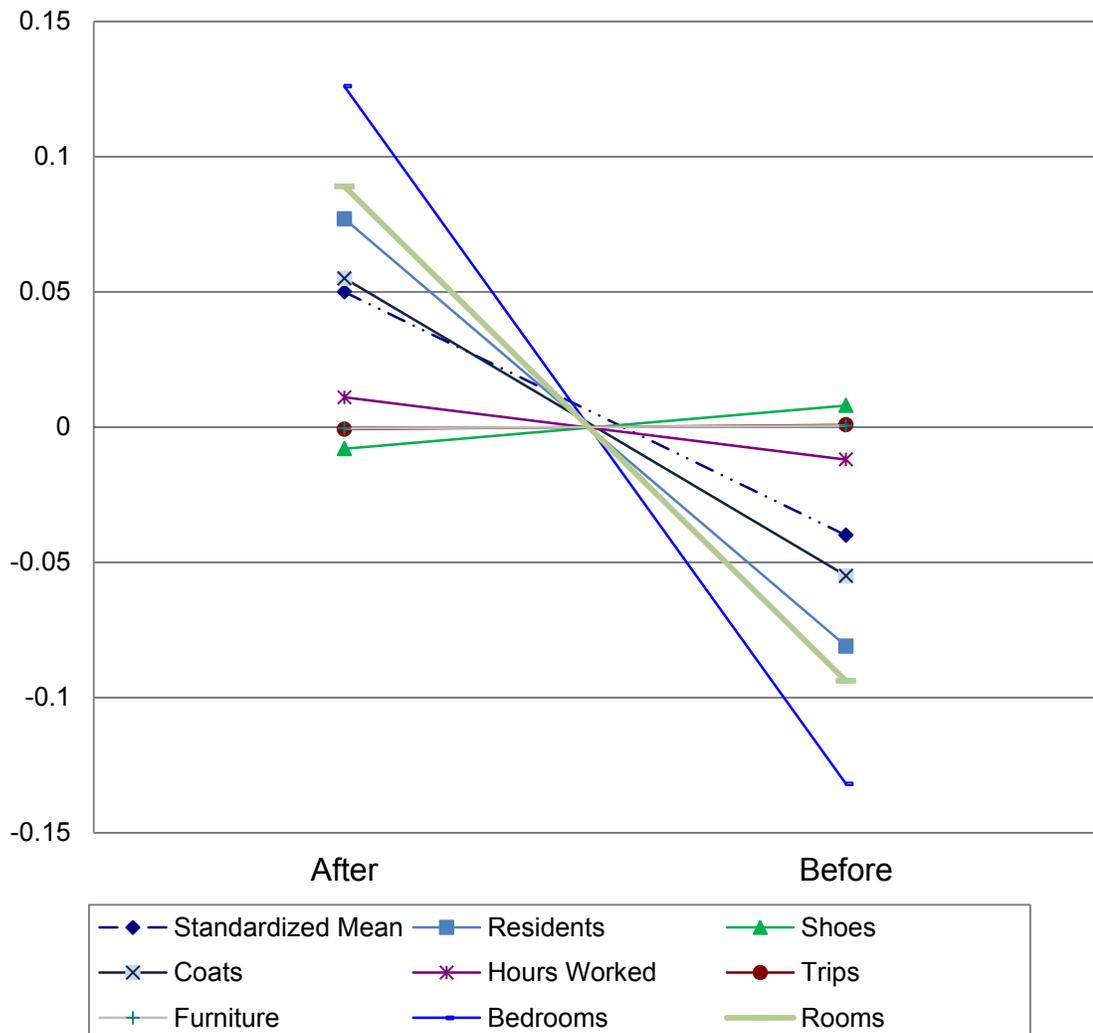


Figure 2.2: Standardized Mean Response to Items as a Function of the Order of Presenting Clarifying Instructions

was higher (positive) when the clarifying instructions came after the question and lower (negative) when the instructions came before the question. A 2 x 2 ANOVA on the combined means revealed a significant main effect for order ($F(1, 577) = 5.20, p < .05$). The 2x2 ANOVA also revealed no significant main effect for font ($F(1, 577) = 0.0, \text{n.s.}$), and no interaction effect between order and font ($F(1, 577) = 0.02, \text{n.s.}$).

2.3.2 Responses to the Follow-up Questions

The second outcome variable was whether respondents answered the *shoe* and *hours worked* items correctly—that is, whether they excluded the subclasses they were to supposed to exclude according to the clarifying instructions. Respondents' answers to the follow-up questions were coded and classified as correct or incorrect. For example, respondents who reported no slippers, boots, sneakers, or athletic shoes in response to the *shoe* follow-up question were coded as consistent with the clarifying information. If, on the other hand, respondents reported such footwear as shoes in the follow-up question, then they were coded as inconsistent with the clarifying information. If respondents reported not owning any shoes (or not having worked) in response to the original questions, the follow-up questions were coded as consistent with the clarifying information. This was done because the questions were worded such that zero was a valid response.

In the multiple question method, the series of subsequent questions were used to determine the “validity” of respondents' answers to the first question in the

series. Responses to the subsequent questions revealed how inconsistent respondents' answers would have been if these invalid categories were not later subtracted out. Thus, if a respondent did not include boots, sneakers, or athletic shoes in the first question of the series according to their response to the next question in the series, the validity check was coded as consistent with the clarifying information. If, on the other hand, such footwear was reported as shoes, then the validity check was coded as inconsistent with the clarifying information.

As shown in Table 2.4, significantly more respondents correctly excluded the relevant subclasses when clarifying instructions were provided in the *shoe* item. About 7% of the respondents excluded the correct subclasses from their responses in the first question of the multiple question method and 1% of the respondents excluded the correct subclasses in the questions with the no clarifying instructions, respectively. This rose to nearly 50% of the respondents excluding the correct subclasses when the questions provided clarifying instructions. As can be seen in Table 2.4, these differences were significant across the three main conditions. Taken together, these results suggest that few respondents reported "correctly" in the absence of clarification, but nearly half of the respondents reported correctly when clarification was provided. Thus, people's answers to the *shoe* question were correct nearly half the time in the presence of clarifying instructions.

For the hours worked item, about 55% of the employed respondents excluded the correct subclasses in the first question of the multiple questions method and even more, 72% of the employed respondents, excluded the correct subclasses in the no

clarifying instruction condition.⁵ Nearly 79% of the employed respondents excluded the correct subclasses in the clarifying instructions condition. Although these percentages were significantly different across all three groups, these percentages were not significantly different between the no clarifying instruction and clarifying instructions group ($\chi^2(1, N = 527) = 2.32, n.s.$). Thus, the *hours worked* question did not display the same pattern as the *shoes* question. The results seem to suggest that the employed respondents had a greater tendency to interpret the question as intended and to answer correctly, despite the absence of clarifying instructions. Thus, the addition of clarifying instructions did not improve their interpretations.

⁵People who were not employed were excluded from the results presented here. However, the overall pattern looked similar when the not employed were included.

Table 2.4: Percentage of “Valid” Responses by Method of Presenting Clarifying Instructions for the Follow-up Items

Follow-up Item	First Question of Multiple Questions		Clarifying Instructions				Main Conditions χ^2	After vs. Before χ^2	
	Percent Valid (n)	Percent Valid (n)	Total	After		Before			
				Percent Valid (n)	Percent Valid (n)	Percent Valid (n)			Percent Valid (n)
Shoes	6.6 (151)	1.1 (176)	49.6 (575)	48.0 (296)	51.3 (279)	199.9***	0.5 n.s.		
Hrs Worked	55.2 (105)	72.2 (115)	78.8 (412)	85.2 (291)	75.6 (205)	24.3***	2.6 n.s.		

Note: People who were not employed were excluded from the hours worked item. Chi squared values for the three main conditions are based on 2 degrees of freedom. Chi squared values for the after vs. before comparison are based on 1 degree of freedom from a 2 x 2 logistic regression. *** $p < .001$, n.s. denotes not significant.

I also compared the four versions of the questions that included clarifying instructions (Table 2.4), using a 2 x 2 logistic regression analysis. The percentage of respondents that provided valid responses to the *shoe* question did not differ significantly by whether the clarifying instructions came before or after the question, although the percentage moved in the predicted direction. Forty-eight percent of the respondents provided valid responses when the clarifying instructions came after the question compared to 51% who provided valid responses when the clarifying instructions came before. Similarly, the percentage of respondents who provided valid responses in the *hours worked* items did not differ significantly by order. In this case, however, the percentage moved in the opposite from the expected direction: 82% of the respondents provided valid responses when the clarifying instructions came after the question compared to a little over 75% when the clarifying instructions came before.

The 2x2 logistic regression analysis also revealed that the percentage of respondents that provided valid responses for either question did not differ significantly by font [*shoes*, $\chi^2(1, N = 571) = 0.3$, n.s. and *hours worked*, $\chi^2(1, N = 408) = 0.4$, n.s.]. Nor was there an interaction between order and font for either item [(*shoes*, $\chi^2(1, N = 571) = 0.07$, n.s. and *hours worked*, $\chi^2(1, N = 408) = 0.03$, n.s.).

2.3.3 Response Times

The third outcome variable was the amount of time it took respondents to read and answer the questions with clarifying instructions. I focused on these four

versions of the items because they include the same number of words. As often occurs in reaction time studies, there were outliers in the data. The slowest one percent of the times for each individual item were removed.⁶

The results here appear clear. As shown in Table 2.5, respondents spent significantly less time on a question when the clarifying instructions came after the question than when the instructions came before. This difference occurred for all eight items. As shown in the last row of Table 2.5, taken together, respondents spent significantly less time, nearly 45 seconds less, reading and processing the clarifying instructions when they came after the question than when they came before. This averages out to about 5 seconds less per item.

There was not a significant main effect in response times for the font ($F(1, 578) = 0.93$, n.s.), nor was there an interaction effect between order and font ($F(1, 578) = 0.54$, n.s.).

I also looked at response times across all three of the main conditions. A one-way ANOVA showed that the effect of time was significantly different by condition ($F(2, 910) = 155.64$, $p < .001$). It took significantly more time to answer the eight questions when they were broken down into multiple questions (335.7 seconds) than it took to answer the eight questions with clarifying instructions (217.7 seconds) or

⁶The cutoffs and the number of values removed for each item was: residents > 95 seconds (6 values); shoes > 253 seconds (6 values); coats > 281 seconds (6 values); hours worked > 446 seconds (6 values); trips > 170 seconds (6 values); furniture > 93 seconds (6 Values); bedrooms > 150 seconds (6 values); rooms > 145 seconds (6 values). For example, 99% of the response times fell between 1.75 and 170.8 seconds for the trip question, but one value was as high as 22085.3 seconds. The conclusions are similar if I used a 90 second cutoff for all eight items.

Table 2.5: Mean Response Time in Seconds (and Sample Sizes) by Order of Presenting Clarifying Instructions and Item

Item	Order				<i>F</i> - statistic
	Clarifying Instructions After		Clarifying Instructions Before		
	Mean	(n)	Mean	(n)	
Residents	24.3	(296)	29.0	(280)	13.5 ***
Shoes	26.7	(297)	34.4	(279)	17.6 ***
Coats	30.3	(298)	36.7	(278)	7.5 **
Hours Worked	25.9	(298)	35.7	(278)	22.4 ***
Trips	25.6	(297)	29.4	(279)	6.1 **
Furniture	18.7	(298)	23.3	(278)	15.7 ***
Bedrooms	15.8	(299)	20.5	(277)	25.9 ***
Rooms	29.7	(299)	34.1	(277)	6.9 **
Total	195.6	(300)	240.1	(282)	27.8***

Note: Slowest one percent of times for each individual item were removed. *F*-statistics for the individual items are from 2-way ANOVAS; all are based on 1 degree of freedom. *** $p < .001$; ** $p < .01$

the eight questions without clarifying instructions (139.1 seconds)

2.4 Discussion

My study addressed several research questions. First, did providing clarifying instructions improve respondents' comprehension of the concepts? The results suggest that providing clarifying instructions worked, though not perfectly. Respondents did seem to correctly restrict their interpretations of potentially ambiguous or vague concepts, and to provide lower responses in the presence of instructions designed to lower their answers. Conrad, Couper, Tourangeau, & Peytchev (2006) have shown that respondents are more likely to pay attention to clarification when it is useful or surprising. For example, telling respondents that French fries and potato chips are vegetables appears to be surprising and useful. Although the clarifying instructions included in my experiment were not specifically written to be useful or surprising, perhaps they came across that way. For example, when reporting the number of shoes they owned, perhaps respondents found it surprising to exclude "sneakers" from their responses, and the clarifying instructions were particularly effective as a result.

Though these had an effect, the clarifying instructions appeared to be only 60% as effective as asking multiple questions. Providing clarification in the residence question reduced the mean response by 20% relative to the the no clarifying control group, but asking multiple questions reduced it by 33%. Further evidence to support the notion that the clarifying instructions did not work perfectly comes

from respondents' answers to the follow-up question to the *shoe* question. If respondents had adhered perfectly to the clarifications in the *shoe* question, we would have expected respondents to have excluded the correct subcategories 100% of the time, but they only appeared to exclude the correct subcategories 50% of the time.

These results are in line with the findings of Tourangeau, Conrad, Arens, Fricker, Lee, & Smith (2006), who used vignettes to examine respondents' understanding of two particular concepts (*residence* and *disability*). The researchers found high levels of classification errors, especially when respondents were provided a technical definition in one of the questions, the *disability* question. Respondents who got a technical definition of *disability* only classified about 50% of the vignettes correctly.⁷

As predicted, some ways of presenting the clarifying instructions in this experiment appeared to be more effective than others. The second research question was whether placing the clarifying instructions before the question would increase their impact on the answers. Since the clarifying instructions had little effect in the *furniture* item to begin with, it is not surprising that the order variable had no effect in this question either. Despite this, there was a difference between the before and after methods over all items that appeared to slightly favor presenting

⁷The technical definition differed from respondents' everyday definitions, and thus, may be more similar to the clarifying instructions used in this experiment. The researchers also provided respondents with everyday definitions. They found that respondents' abilities to classify the vignettes with definitions that were modeled after their everyday sense of a word were significantly better (in this case, respondents classified about 75% of the vignettes correctly for the disability question).

the clarification before the question.

Although previous research has demonstrated that respondents sometimes interrupt clarifying instructions that come after a question in the aural channel (e.g., Houtkoop-Steenstra, 2002; Oksenberg, Cannell, & Kalton, 1991; Van der Zouwen & Dijkstra, 2002), it was not clear whether respondents would pay less attention to the clarifying instructions that came after a question presented visually. The fact that the standardized mean responses were significantly lower when the clarifying instructions came before the question than afterwards suggests that respondents do pay a little less attention to clarifying instructions that follows a conversational endpoint (the end of a question) and to answer prematurely, even with questions presented visually. Additional evidence to suggest that respondents pay less attention to the clarifying instructions when they are placed after the question comes from the analysis of response times. Respondents devoted significantly less time on the study questions when the clarifying instructions followed the questions than when they preceded them. The results of the follow-up question analysis were not so clearly supportive, however; responses to the before and after methods did not significantly differ. It may have been that there was not enough statistical power to detect a difference here, since the effect sizes were small and the sample sizes for the follow-up questions were relatively small. In hindsight, it seems unfortunate that I chose the *hours worked* item as one of the questions for this part of the study, since it was an item that consistently behaved differently from the other seven questions (and moved in the opposite direction of expectations).

A third research question was whether it was more effective to put the clari-

fying instruction in the same font as the question or to use italics. There were two conflicting theories under examination here. One theory suggested that respondents would be more likely to view the clarifying instruction as on the critical path when it was placed in the same font as the question (Jenkins & Dillman, 1997). The other theory suggested that respondents would be more likely to process the clarifying instructions when they were differentiated from the question text (McCarthy & Mothersbaugh, 2002). According to all three analyses (mean responses, response times, and follow-up questions), respondents were no more likely to process the clarifying instructions when they were in the same font as the question than when they were not.

The final question was whether reducing the complexity of the task by asking multiple questions improved comprehension even further. As predicted, incorporating the clarifying instructions into a series of questions did lower the mean response more than presenting instructions did. This approach may have forced respondents to pay more attention to the clarifying instructions. In contrast, respondents could still view the clarifying instructions as separate from the question when they were presented before or after the question, and could still skip over the clarifying instructions to get to either the question or the response options in these versions.

Another reason for the lowered mean response in the multiple question method might be that respondents need not hold as much information in memory. The questions were shorter, so they required keeping less verbal information in memory. Also, respondents need only report the sub-quantities sequentially under this method, which meant they need not need to hold these quantities in memory nor

did they need to perform any mental arithmetic with them.

Getting respondents to understand complex concepts does appear to take more time. It took respondents about 42 seconds, on average, to read and answer a series of questions in the multiple question condition. In comparison, it only took about 27 seconds to read and respond to one of the questions with clarifying instructions. In addition, a few respondents reported negative responses in the multiple question condition that needed to be removed. This problem can probably be mitigated by letting respondents know they had made such a mistake electronically. There was no evidence from this study that respondents had a greater tendency to satisfice in the multiple question condition, that is, to provide zero in response to the first question in the series so as not to have to answer further questions. In addition, there was no evidence that they went back and changed their response to the first question in the series as a result of reading and answering later questions in the series. The means of the first question in the series were not significantly different from the means of the no clarifying questions. All in all, it appears that the advice drawn from theory that it is better to incorporate the clarifying instructions into a series of shorter questions has merit (Conrad & Couper, 2004; Conrad & Schober, 2000; Couper, 2008, p. 289; Fowler, 1995, pp. 13-20; Jenkins & Dillman, 1997; Schaeffer & Presser, 2003; Sudman, Bradburn, & Schwarz, 1996, p. 31; Suessbrick, Schober, & Conrad, 2000; Tourangeau, Rips, & Rasinski, 2000, pp. 38-40, 61).

2.5 Conclusions and Future Research

This chapter provides evidence from a probability sample responding via the Internet that respondents' interpretations of ambiguous and/or vague concepts can be improved by providing clarifying instruction. As hypothesized, it also seems that some methods are better than others. Strategically placing information before the question appears to be beneficial. There is modest evidence to suggest that placing clarifying instructions before the question increases the likelihood that it will be seen as being on the critical path and processed. Thus, from a practical perspective, if survey practitioners are restricted to asking one question only, they should consider placing the clarifying instructions before the question rather than after it. The practice of attempting to highlight any clarifying instructions by putting them in italics is not supported by this research. However, in line with many researchers' recommendations, the practice of breaking the questions into a series of questions that incorporate the clarifying instructions appears to be the most effective approach. Although readers may potentially be better equipped to understand longer questions than listeners, it would seem that respondents in the visual channel benefit from shorter questions as well.

This study was conducted in the visual channel only – that is, via a Web survey. Future research should be aimed at extending this research to the aural channel to compare the clarifying methods across channels. The research presented in this chapter was designed such that lower responses were presumed to be more accurate or valid. Future research should verify this presumption. The study items could

be examined in different settings, perhaps providing respondents with scenerios on which to base their answers, or asking a set of items for which validity data are available. Finally, the research described here did not focus upon the use of examples, which have also been shown to aid respondents' understanding. Future research should examine methods of presenting examples in surveys that will improve their use as well.

Chapter 3

Crossing Clarifying Methods and Sensory Channels

3.1 Introduction

Past research has shown that ambiguous or vague concepts in survey questions can be clarified through the use of definitions, instructions or examples (e.g., Conrad, Schober, & Coiner, 2007; Tourangeau, Conrad, Couper, & Ye, 2010), but respondents do not necessarily attend to these clarifications (e.g., Tourangeau, Conrad, Arens, Fricker, Lee, & Smith, 2006). The aim of this research is to investigate how to present clarifying information so that respondents read it carefully. A key issue is whether questions presented aurally make different demands on comprehension from those presented visually. The answer to this question is not clear because channel is typically confounded with the presence of an interviewer in many mode comparison studies (e.g., Bishop, Hippler, Schwarz, & Strack, 1988; Chang & Krosnick, 2009; Christian, Dillman, & Smyth, 2007; Fricker, Galesic, Tourangeau, & Yan, 2005; Heerwegh & Loosveldt, 2008; Smyth, Christian, & Dillman, 2008).

A recent Web experiment found that respondents' interpretations of terms in survey questions could be changed by providing clarifying instructions (see Chapter 2). For example, that study found that instructing respondents to exclude boots, sneakers, athletic shoes, and bedroom slippers reduced the average number of shoes reported by respondents. Some methods of providing clarifying instructions ap-

peared more effective in that experiment than others. Specifically, incorporating the clarifying instructions into a series of questions appeared more effective than providing instructions either before or after the question. In addition, that study found that placing the clarifying instructions before the question was somewhat more effective than placing them after the question. However, these findings are limited to a Web survey. It is unclear whether these findings extend to the aural channel as well and whether there might be an interaction between the method of providing instructions and the channel of presentation.

There are reasons to think that the channel in which the instructions are provided will influence how much attention they receive. Both readers and listeners may need clarifying instructions to understand the concepts as intended, but features of visual presentation may make it easier for respondents to comprehend the instructions compared to aural presentation, especially when the instructions are complex (e.g., Clark & Brennan, 1991; Just & Carpenter, 1980; Osada, 2004; Rayner & Clifton, 2009; Schwarz, Strack, Hippler, & Bishop, 1991a). These features of visual modes include respondent control over the pace of input in the visual channel, the ability to preview (that is, to see that something is present without necessarily reading it) and review input, and the ability to repair any errors in private. Evidence from the persuasion literature and studies in listenability have shown that written (vs. videotaped or audiotaped) messages enhance comprehension when the material is complex (e.g., Chaiken & Eagly, 1983; Dickens, Harwood, & Carter, 1955). A final difference between visual and auditory presentation that may contribute to differential attention is that text is supplemented by visual cues, such

as color, shape, and location, whereas speech is supplemented by aural cues, such as stressed words and variations in pace (Redline & Dillman, 2002; Tourangeau, Couper, & Conrad, 2004; Wennerstrom, 2001, p. 4). These cues may differentially affect respondents' attention.

In this chapter, I report on an experiment that examines the effects of providing clarifying instructions visually and aurally. The goal was to gain a better understanding of how to improve survey questions across modes. This issue is especially relevant given the current debate over how to design questions for mixed-mode surveys (De Leeuw, 2005; Dillman, 2000; Martin, Hunter Childs, DeMaio, Hill, Reiser, Gerber, Styles, & Dillman, 2007a). I predicted that respondents would understand questions that included clarifying instructions better in a Web survey (where instruction is provided visually) than an Interactive Voice Response (IVR) survey (where the instruction is presented aurally). However, I also thought that asking a series of questions rather than one question would reduce any channel differences and lead to better answers in both channels. Multiple questions would, I thought, simplify the task for respondents and allow them greater time on task in both channels.

3.2 Methods

Respondents were contacted by telephone interviewers from Abt SRBI and asked to participate in a study about health practices and lifestyles on behalf of the University of Maryland. After answering a few background questions and agreeing

to participate in the study, respondents were assigned to one of two modes of data collection – Web or IVR – that primarily differ in their sensory channel. IVR was chosen as the aural mode because it is also a self-administered mode of data collection. Abt SRBI carried out the data collection in both modes.

3.2.1 Sampling and Data Collection

A list-assisted landline Random Digit Dial (RDD) sample was selected. Cases that were successfully contacted and that screened in ($n = 1,304$) were assigned to a mode of data collection. The eligible population consisted of adults living in residences with landline telephones and Internet access; cell telephones were excluded from the sample. Because interviews were conducted in English, it is further restricted to the English speakers. In a list-assisted sample, random numbers are appended to randomly selected eight-digit blocks (xxx-yyy-zz) associated with one or more listed numbers to form “lines” or potential telephone numbers. Survey Sampling International provided the sample lines. Twenty-two thousand (22,183) potential landlines were dialed.

The survey began with a short screening interview administered in Computer Assisted Telephone Interviewing (CATI). The screener portion of the interview asked respondents about their access to the Internet and included a few demographic questions (e.g., age, education, sex, and race). One adult was selected in each eligible household using the “last birthday” method. Respondents without Internet access were screened out. The incidence for Internet access was 75% among those

completing the screener. Respondents with Internet access were initially randomly assigned (with equal probability) to the Web or IVR conditions. To achieve a roughly equal number of completed interviews in each channel, after half of the interviews were conducted, a smaller fraction of cases were assigned to IVR and a greater fraction was assigned to Web.

The screening interviews were conducted from February 25 to April 8, 2010. The response rate to the screening interview was nearly 22% – 1,780 screeners were completed out of an estimated 8,199 working residential numbers (AAPOR RR1). The 1,304 screener respondents with Internet access were assigned to a mode of data collection; 475 were assigned to the IVR mode and 829 were assigned to the Web mode.

3.2.1.1 IVR Cases

The IVR interviews were completed from February 25 through April 3, 2010. Screener respondents assigned to the IVR mode were offered \$5.00 to complete the IVR survey. The incentive check was mailed to respondents after they completed the survey. Respondents who agreed to participate were transferred by the interviewer to the IVR system at the conclusion of the telephone screening interview. Nearly 43% (204/475, AAPOR RR1) of the respondents assigned to the IVR mode completed the survey. Twenty two percent refused to be transferred to the IVR system, and 20% started but did not complete the survey (most of these were transferred but did not start the IVR survey). The average length for completed IVR surveys was

12.1 minutes.

A few days after screening interviews had begun (March 1, 2010), a slight change was made to the wording of the screening interview to reduce the number of respondents who started the IVR survey, but did not complete it. These changes included reminding respondents about the incentive (\$5.00), the sponsor (the University of Maryland), and asking that they have patience with the IVR system.

3.2.1.2 Web Cases

Web surveys were completed from February 26 through April 15, 2010. Screener respondents assigned to the Web were offered a \$10 incentive – an electronic debit card sent by email – for completing the Web survey. At the conclusion of the telephone screener, all respondents who agreed to do the Web survey were immediately sent an invitation email with the URL for the questionnaire and a unique login ID.¹

Respondents with email addresses who did not complete the Web survey were sent a reminder email after three days and a second reminder email after seven days. A final email reminder was sent during the final week of the field period. One telephone reminder was also made to non-respondents.

About 24% (203/829, AAPOR RR1) of the respondents assigned to the Web mode completed the survey. Thirty-four percent of the respondents refused to participate in the Web survey; 41% agreed to do the survey but did not start it, and 1% got part way through the Web survey but did not complete it. The average length

¹Thirty-five respondents who declined or did not have email were read the URL and login ID. In three instances, respondents were faxed the URL and their unique login ID.

for completed Web surveys was 12.3 minutes.

3.2.2 Study Questions and Clarifying Methods

Seven of the eight study questions from the previous experiment (described in Chapter 2) were also used in this experiment. Since most people had not reported purchasing any furniture in the previous experiment, the furniture question was dropped and replaced with a question about doctor visits. A power analysis suggested that adequate power for this study required a total of ten questions. Thus, two additional questions about the number of telephone calls the respondent made or received and the number of emails sent by the respondent were developed. The ten study questions asked in this experiment were:

1. How many people are currently living or staying at your home address?
2. How many pairs of shoes do you own?
3. How many coats and jackets do you own?
4. In the past week, how many telephone calls did you make or receive?
5. Last week, how many hours, if any, did you work for either pay or profit?
6. In the past year, how many times, if any, were you away from home on a trip?
7. In the past week, how many email messages, if any, have you written?
8. In the past year, how many times, if any, have you seen or talked to a medical doctor?

9. How many bedrooms are in your house, apartment, or mobile home?
10. How many other separate rooms are in your house, apartment, or mobile home?

Most of these questions were modeled on items from ongoing federal surveys. For example, question 1 is similar to the initial item on the decennial census form. Each of the questions included clarifying instructions. The clarifications were meant to restrict respondents' interpretations of the question's core concept (see Appendix A.2 for the complete questionnaire, including all clarifying instructions). For example, the clarifying instruction for the shoe question (2 in the list above) was:

For the purposes of this question, do not include boots, sneakers, athletic shoes, or bedroom slippers. Include sandals, other casual shoes, and dress shoes. If you do not own a pair of shoes (as we have defined them), enter "0."

As shown in Table 3.1, the study questions were administered with the clarifying instructions coming before or after the main question text or incorporated in several questions.

Initially, respondents were randomly assigned to the three clarifying instruction conditions with an equal probability of assignment to each method. After approximately three-fourths of the interviews were done, these probabilities were changed to achieve more nearly equal sample sizes (Wei, 1978).

The two additional questions used in the previous experiment to better understand respondents' answers to the *shoe* and *hours worked* questions were also

Table 3.1: Examples of Alternative Clarifying Methods

Method 1	Method 2	Method 3
<p data-bbox="363 569 639 625">2A. The next question is about your footwear.</p> <p data-bbox="363 659 656 1016">How many pairs of shoes do you own? For the purposes of this question, do not include boots, sneakers, athletic shoes, or bedroom slippers. Include sandals, other casual shoes, and dress shoes. If you do not own a pair of shoes (as we have defined them), enter “0.”</p> <p data-bbox="363 1052 634 1079">Number of pairs of shoes</p> <div data-bbox="412 1087 532 1161" style="border: 1px solid black; width: 74px; height: 35px; margin-left: 20px;"></div>	<p data-bbox="691 569 964 625">2B. The next question is about your footwear.</p> <p data-bbox="691 659 984 989">For the purposes of this question, do not include boots, sneakers, athletic shoes, or bedroom slippers. Include sandals, other casual shoes, and dress shoes. If you do not own a pair of shoes (as we have defined them), enter “0.” How many pairs of shoes do you own?</p> <p data-bbox="691 1024 959 1052">Number of pairs of shoes</p> <div data-bbox="740 1060 860 1134" style="border: 1px solid black; width: 74px; height: 35px; margin-left: 20px;"></div>	<p data-bbox="1024 569 1313 625">2C1. The next question is about your footwear.</p> <p data-bbox="1024 659 1346 716">How many pairs of shoes do you own?</p> <p data-bbox="1024 751 1305 779">Number of pairs of shoes</p> <div data-bbox="1102 779 1222 852" style="border: 1px solid black; width: 74px; height: 35px; margin-left: 20px;"></div> <p data-bbox="1024 873 1175 900">[If 2C1 > 0:]</p> <p data-bbox="1024 905 1360 1079">2C2. When you reported the pairs of shoes that you own, how many pairs of boots, sneakers, athletic shoes, or bedroom slippers, if any were included?</p> <p data-bbox="1024 1115 1297 1142">Number of pairs of shoes</p> <div data-bbox="1096 1142 1216 1230" style="border: 1px solid black; width: 74px; height: 42px; margin-left: 20px;"></div> <p data-bbox="1024 1268 1365 1295">[If 2C1=2C2, do not ask 2C3.]</p> <p data-bbox="1024 1299 1360 1446">2C3. When you reported the pairs of shoes that you own, how many pairs of sandals, other casual shoes, or dress shoes, if any, were included?</p> <p data-bbox="1024 1482 1305 1509">Number of pairs of shoes</p> <div data-bbox="1096 1514 1216 1593" style="border: 1px solid black; width: 74px; height: 38px; margin-left: 20px;"></div>

included in this one. The purpose of these questions was to uncover the subcategories respondents had included in their responses to the *shoe* and *hours worked* questions to determine the accuracy of their responses to those questions. In other words, had they included boots, sneakers, or bedroom slippers in the *shoe* question, despite being instructed not to?

A respondent received all of the questions in the same clarification method and mode combination. Twenty-six questions were asked in total. The ten study and two follow-up questions came as items 1, 6, 7, 8, 12, 13, 14, 18, 19, 20, 24, and 25 in the questionnaire. To limit any carryover effects, the study questions were separated by buffer questions (often requiring only yes or no responses). An example of one of the buffer questions is “Many people say they have less time these days to do volunteer work. What about you, were you able to devote any time to volunteer work in the last 12 months?”

Both the Web and the IVR instrument discouraged item non-response through the use of a “soft” prompt, which requested respondents to provide their best answer when they left an item blank. The IVR instrument was programmed to allow respondents to interrupt the reading of a question and to answer. To maintain comparability between the one question methods with the multiple question method in the visual channel, all of the questions for a given item (such as shoes) were displayed on one Web screen.

3.3 Results

3.3.1 Effects of Nonresponse

Table 3.2 shows the dispositions of those who agreed to participate in the survey and who actually accessed either computer system. As can be seen, a greater percentage of those who accessed the IVR system either broke off during the course of answering the questions (breakoffs) or quit before answering any questions at all (nonstarters) compared to those that accessed the Web system.² However, there were no differences in the percentage of respondents who broke off or did not start the survey by clarifying method.

Table 3.3 shows that there were no significant differences in the demographic characteristics or the reported behaviors of those who answered the survey in the different mode conditions, with one exception. Significantly more employed persons responded via the Web (71.6 %) than the IVR (61.4%).

3.3.2 Responses to the Study Items

Answers to the study questions were numeric and required some editing before they could be analyzed. In the multiple question method, responses to each item were derived by subtracting respondents' answers to the subsequent questions from their answers to the first question in the series. Negative values that resulted from this calculation were set to missing. Values that were above the upper one percentile

²Fifty-six of the cases were assigned to mode of data collection, but dropped out before being assigned to one of the methods of clarification.

Table 3.2: Percentage of Respondents Who Accessed Either Computer System (and Sample Sizes) by Outcome, Mode of Data Collection, and Method of Clarification

Condition	Completes	Breakoffs	Nonstarters	(n)	χ^2 Statistic
Mode					
IVR	55.7 %	25.1 %	19.1 %	(366)	(2, N = 573) = 115.2 ***
Web	98.0	1.5	0.5	(207)	
Method					
After	80.7	16.8	2.4	(166)	(1, N = 517) = 1.8 n.s.
Before	79.7	16.9	3.5	(172)	
Multiple	76.0	21.2	2.8	(179)	

Note: *** $p < .001$; n. s. denotes not significant

for each item were also removed.³

Out-of-range values are defined as negative values and outliers (values greater than the upper one percentile). As shown in Table 3.4, there was a highly significant main effect for mode in the percentage of out-of-range values – 5.1 % of the responses

³The number of negative values set to missing for each item were: residents (one value); shoes (five values); coats (nine values); telephone calls (three values); hours worked (four values); trips (three values); emails (one value); doctor visits (one value); bedrooms (zero values); rooms (17 values). The cutoff and number of high values removed for each item were: residents ≥ 46 (six values); shoes ≥ 80 (five values); coats ≥ 25 (five values); telephone calls ≥ 125 , (three values); hours worked ≥ 82 (five values); trips ≥ 25 (four values); emails ≥ 250 (four values); doctor visits ≥ 50 (five values); bedrooms ≥ 6 (two values); rooms ≥ 12 (four values). Removing the negative and extreme values did not change the overall outcomes of the analyses.

Table 3.3: Percentage of Respondents (and Sample Sizes) by Demographic Characteristics or Reported Behaviors and Mode of Data Collection

Characteristic	Web		IVR		χ^2 Statistic
	Percent	(n)	Percent	(n)	
Male	36.9 %	(206)	43.8 %	(290)	(1, N = 496) = 2.37, n.s.
Female	63.1		56.2		
Black	7.4	(204)	7.7	(285)	(2, N = 489) = 5.25, n.s.
White	83.8		88.4		
Other	8.8		3.9		
Hispanic	3.9	(205)	7.2	(290)	(1, N = 495) = 2.43, n.s.
Not Hispanic	96.1		92.8		
High school or less	15.5	(206)	17.7	(288)	(5, N = 494) = 7.00, n.s.
Some college	20.9		24.7		
College degree	63.6		57.6		
18 - 34 years	14.7	(198)	13.0	(284)	(3, N = 482) = 2.07, n.s.
35 - 44 years	19.2		16.6		
45 - 54 years	31.3		29.2		
55 years and older	34.9		41.2		
Volunteered	58.3	(202)	59.8	(246)	(1, N = 450) < 1, n.s.
Did not volunteer	41.7		40.2		
Would serve jury duty	63.7	(204)	55.8	(242)	(1, N = 446) = 2.89, n.s.
Not serve jury duty	36.3		44.2		
Made contributions	84.3	(204)	85.6	(243)	(1, N = 447) < 1, n.s.
Did not contribute	15.7		14.4		
Employed	71.6	(201)	61.4	(228)	(1, N = 429) = 5.00 *
Not employed	28.4		38.6		
Traveled to work by:					(2, N = 251) = 1.00, n.s.
Personal vehicle	88.6	(131)	88.3	(120)	
Public transportation	5.3		3.3		
Other	6.1		8.3		
Quality of life in community:					(3, N = 412) = 5.19, n.s.
Excellent	36.5	(203)	35.4	(209)	
Good	49.8		44.0		
Fair	12.8		16.3		
Poor	1.0		4.3		
Economic conditions in community:					(3, N = 413) = 3.64, n.s.
Excellent	6.4	(203)	10.0	(210)	
Good	44.3		39.5		
Fair	38.4		35.7		
Poor	10.8		14.8		
Joined organization	26.1	(203)	32.9	(207)	(1, N = 410) = 2.24, n.s.
Did not join organization	73.9		67.1		

Note: * $p < .05$; n.s. denotes not significant

were out-of-range in the IVR compared to 1.2 % in the Web. There was also a significant main effect for clarifying instruction methods. The multiple-question method yielded the largest percentage of out-of-range values, 5.1 %, compared to the versions with clarifying instructions before or after the question, which yielded 2.9 % and 2.3 %, respectively.

Table 3.5 presents the individual means for each of the items by method of presenting the clarifying instructions and data collection mode, and Table 3.6 presents the results of the individual two-way ANOVAs. Initially, these results were stratified by changes to the screening questionnaire (before March 1 versus March 1 and after), but as this was not a theoretically motivated variable and there were no significant differences in respondents' answers between these time periods, neither individually nor overall, this variable was dropped from further analysis. All analyses were also stratified by a second time period variable to control for differential assignment to mode (February 25 - March 11, 2010 versus March 12 - March 25, 2010 versus March 26 - April 4, 2010). Introduction of this second time period variable into the analyses also had little effect on the conclusions, either for the individual items or overall.

Table 3.5 presents means for each item and 3.6 presents results from ANOVAs (from models including just the two experimental variables). I expected the means to be highest when the clarifying instructions came after the question text, lowest when the item was broken into several questions embodying the restrictions in the instructions, and intermediate when the clarifying instructions came before the question text. As can be seen, four of the ten items (*residents*, *coats*, *bedrooms*, and

Table 3.4: Percentage of Out-of-Range Values and Percentage of Item Missing (and Sample Sizes) by Mode of Data Collection and Method of Presenting Clarifying Instructions

		Clarifying Method								
Error	Mode	Clarifying Instructions After		Clarifying Instructions Before		Multiple Questions		Total	χ^2 Statistic	
		%	(n)	%	(n)	%	(n)	%		(n)
Out of Range	IVR	3.7	(93)	4.5	(97)	6.8	(106)	5.1	(296)	Mode 44.5***
	Web	0.6	(69)	0.6	(69)	2.4	(68)	1.2	(206)	Method 21.0***
	Total	2.3	(162)	2.9	(166)	5.1	(174)	3.5	(502)	Interaction 1.4 n.s.
Missing	IVR	1.6	(67)	1.0	(69)	0.9	(68)	1.1	(204)	Mode < 1 n.s.
	Web	0.0	(67)	0.0	(68)	0.0	(68)	0.0	(203)	Method < 1 n.s.
	Total	0.8	(134)	0.5	(137)	0.4	(136)	0.6	(407)	Interaction < 1 n.s.

Note: Out-of-range values are defined as negative values and values above the upper one percentile. Percent out-of-range are computed with those who provided a response (the complete and break-off cases). Percent missing are computed with the complete cases only. Chi-squared values are based on 1 degree of freedom for the mode comparison and 2 degrees of freedom for the method and interaction comparisons. *** $p < .001$, n. s. denotes not significant

Table 3.5: Mean Response (and Sample Sizes) by Method of Presenting Clarifying Instructions, Mode of Data Collection, and Item

Item	Mode	Clarifying Method						Total	
		Clarifying Instructions After		Clarifying Instructions Before		Multiple Questions			
		Mean	(n)	Mean	(n)	Mean	(n)	Mean	(n)
Residents	IVR	2.6	(88)	2.4	(93)	2.2	(102)	2.4	(283)
	Web	2.5	(69)	2.2	(69)	2.0	(68)	2.2	(206)
	Total	2.6	(157)	2.3	(162)	2.1	(170)	2.3	(489)
Shoes	IVR	11.6	(83)	9.1	(88)	8.6	(91)	9.7	(262)
	Web	10.5	(68)	12.4	(68)	9.7	(67)	10.9	(203)
	Total	11.1	(151)	10.5	(156)	9.1	(158)	10.2	(465)
Coats	IVR	4.9	(80)	3.9	(77)	3.3	(87)	4.0	(244)
	Web	4.4	(68)	4.3	(68)	3.6	(62)	4.1	(198)
	Total	4.7	(148)	4.1	(145)	3.3	(149)	4.05	(442)
Telephone Calls	IVR	16.1	(78)	13.5	(75)	15.2	(81)	15.0	(234)
	Web	18.8	(68)	16.9	(68)	17.8	(67)	17.8	(203)
	Total	17.4	(146)	15.1	(143)	16.4	(148)	16.3	(437)
Hours Worked	IVR	16.1	(79)	27.2	(74)	20.4	(77)	21.1	(230)
	Web	26.3	(68)	26.7	(67)	24.2	(68)	25.5	(203)
	Total	20.8	(147)	27.0	(141)	22.1	(145)	23.3	(433)
Trips	IVR	3.0	(73)	2.8	(64)	1.6	(77)	2.4	(214)
	Web	2.2	(67)	1.9	(67)	2.0	(66)	2.0	(200)
	Total	2.6	(140)	2.4	(131)	1.7	(143)	2.2	(414)
Emails	IVR	10.2	(69)	10.0	(68)	12.9	(70)	11.0	(207)
	Web	16.9	(67)	14.9	(67)	13.4	(68)	15.0	(202)
	Total	13.5	(136)	12.4	(135)	13.1	(138)	13.0	(409)
Doctor Visits	IVR	3.8	(69)	3.6	(68)	4.9	(69)	4.1	(206)
	Web	3.8	(67)	3.1	(67)	4.2	(67)	3.7	(201)
	Total	3.8	(136)	3.4	(135)	4.6	(136)	3.9	(407)
Bed-rooms	IVR	2.8	(66)	2.4	(68)	1.7	(70)	2.3	(204)
	Web	2.8	(66)	2.5	(68)	2.1	(68)	2.4	(202)
	Total	2.8	(132)	2.4	(136)	1.9	(138)	2.35	(406)
Rooms	IVR	3.8	(66)	4.2	(65)	2.5	(55)	3.6	(186)
	Web	3.2	(67)	3.6	(68)	2.1	(63)	3.0	(198)
	Total	3.5	(133)	3.9	(133)	2.2	(118)	3.3	(384)

Note: Negative values and values in the upper one percentile for each individual item were removed.

Table 3.6: Individual Statistics (ANOVA) for Main and Interaction Effects, by Item

Items	Main and Interaction Effects	
Residents	Method	$F(2, 483) = 2.97$ *
	Mode	$F(1, 483) = 1.01$ n.s.
	Interaction	$F(2, 483) = 0.07$ n.s.
Shoes	Method	$F(2, 459) = 1.43$ n.s.
	Mode	$F(1, 459) = 1.22$ n.s.
	Interaction	$F(2, 459) = 1.53$ n.s.
Coats	Method	$F(2, 436) = 5.36$ **
	Mode	$F(1, 436) = 0.03$ n.s.
	Interaction	$F(2, 436) = 0.67$ n.s.
Telephone Calls	Method	$F(2, 431) = 0.63$ n.s.
	Mode	$F(1, 431) = 3.16$ †
	Interaction	$F(2, 431) = 0.03$ n.s.
Hours Worked	Method	$F(2, 427) = 3.46$ *
	Mode	$F(1, 427) = 5.13$ *
	Interaction	$F(2, 427) = 2.44$ †
Trips	Method	$F(2, 408) = 2.37$ †
	Mode	$F(1, 408) = 1.48$ n.s.
	Interaction	$F(2, 408) = 1.50$ n.s.
Emails	Method	$F(2, 404) = 0.03$ n.s.
	Mode	$F(1, 404) = 4.66$ *
	Interaction	$F(2, 404) = 0.98$ n.s.
Doctor Visits	Method	$F(2, 401) = 1.66$ n.s.
	Mode	$F(1, 401) = 0.55$ n.s.
	Interaction	$F(2, 401) = 0.09$ n.s.
Bedrooms	Method	$F(2, 400) = 19.56$ ***
	Mode	$F(1, 400) = 1.43$ n.s.
	Interaction	$F(2, 400) = 1.43$ n.s.
Rooms	Method	$F(2, 378) = 20.30$ ***
	Mode	$F(1, 378) = 6.38$ **
	Interaction	$F(2, 378) = 0.12$ n.s.

Note: Negative values and values in the upper one percentile for each individual item were removed. *** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .1$; n.s. denotes not significant

rooms) showed the predicted effect of the method of clarification and the results for one more item was marginally significant (*trips*). The overall trend was also significantly downward across the three main conditions for these items. One more item (*shoes*) also moved downward, although not significantly. The *hours worked* item showed a significant effect for the method of clarification. However, there was a significant reversal between the before and after methods for this item ($F(1, 427) = 5.48, p < .01$).

On average, placing the clarifying instructions before the question reduced the mean response by about 5%, and asking multiple questions reduced the mean response about 15% compared to placing the instructions after the question in both cases. If we drop the *hours worked* question from this analysis because it performed so differently from the other questions, placing the clarifying instructions before the question resulted on average in an 8% reduction in the mean response and asking multiple questions resulted in a 16% reduction, compared to placing the instructions after the question.

The effect for mode was less consistent. Three of the items showed a significant effect for mode (*hours worked*, *emails*, and *rooms*), and one more was marginally significant (*telephone calls*). However, only one of these items, *rooms*, moved in the direction predicted – with lower answers reflecting more careful processing of the clarifying instructions in the Web survey. Three more nonsignificant items (*residents*, *trips*, and *doctor visits*) moved in the direction predicted.

To examine the pattern across all ten items, I standardized the numeric responses for each item (by subtracting the item's overall mean and dividing by its

standard deviation) and used these z-scores to create an overall average for each respondent. The average z-scores were analyzed in a two-way ANOVA.

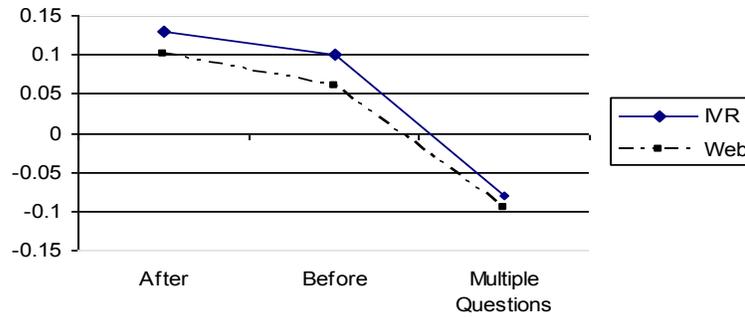
Panel a of Figure 3.1 displays the standardized mean response for each of the clarification methods by mode. A two-way ANOVA shows a significant main effect for clarification method ($F(2, 491) = 4.90, p < .01$). The trend is downward across the three groups. Although the mode difference was in the expected direction (lower for the Web than the IVR), it was not significant ($F(1, 491) < 1$), nor did mode interact with the clarification method ($F(2, 491) < 1$).⁴

The questions that elicited low counts from respondents, such as the number of residents in the household, behaved somewhat differently from those eliciting high counts, such as the number of telephone calls and the number of hours worked. I compared the items with means below the overall mean for the ten items ($M = 8.09$) with those whose means were higher than the overall mean. The six low-count items included *residents*, *coats*, *trips*, *doctor visits*, *bedrooms*, and *rooms*. The four high-count items included *shoes*, *telephone calls*, *hours worked*, and *emails*.

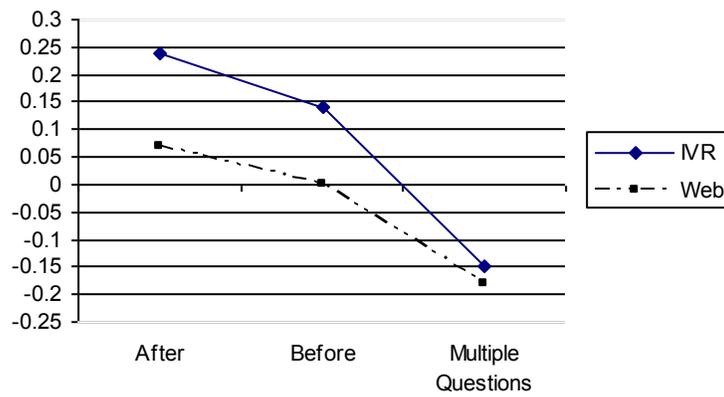
Panel b of Figure 3.1 displays the the standardized mean response for the low-count items by clarification method and mode. Differences between the clarifying methods and the modes look even more pronounced for these items (Panel b) than for all items (Panel a). A two-way ANOVA confirmed this, as the main effect for clarification method was highly significant for the low-count items ($F(2, 489)$

⁴I reran the analyses of variance for the 12 items, this time including employment as a factor. None of the major conclusions changed, and there were no interactions between the experimental variables with employment.

a) All Ten Items



b) Six Low-Count Items



c) Four High-Count Items.

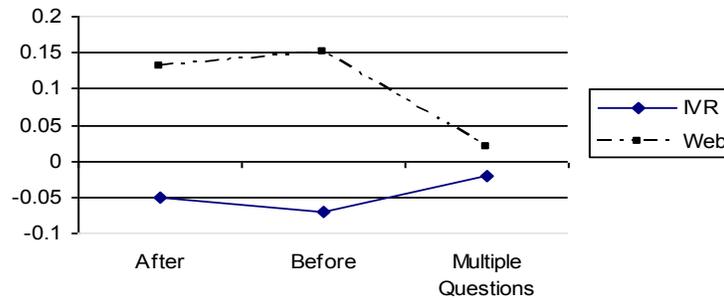


Figure 3.1: Standardized Mean Response Across Items as a Function of Clarifying Method and Channel for All Items (top panel), Items Involving Low Counts (middle panel) and High Counts (bottom panel)

= 9.97, $p < .001$). As before, clarification method trended downward across the two groups ($F(1, 489) = 8.50, p < .01$). However, this time, there was a marginally significant main effect for mode ($F(1, 489) = 2.83, p < .1$) in the predicted direction. A significant interaction effect between clarification method and channel was not evident ($F(2, 489) < 1, \text{n.s.}$).

In contrast to these findings, Panel c of Figure 3.1 displays an entirely different relationship for the high-count items. With these four items, there was no main effect for clarifying method ($F(2, 468) < 1$). In addition, this analysis produced an unexpected finding. There was a significant reversal for the high-count questions by channel, such that high-count questions administered in the aural channel elicited lower means than the same questions administered in the visual channel ($F(1, 468) = 5.61, p < .05$).

To determine whether the pattern exhibited by the low-count items differed significantly from the pattern exhibited by the high-count items, I carried out a repeated measures ANOVA. There were significant interactions between item type and clarification method ($F(2, 466) = 4.95, p < .01$) and item type and mode ($F(1, 466) = 7.89, p < .01$). Taken together, these results confirm that the low-count items seem to behave as hypothesized, but the high-count questions seem to perform differently.

3.3.3 Zero Responses

Respondents may have realized that if they provided a non-zero answer to the initial question of the multiple-question method, they would have to answer a series of follow-up questions. Reporting zero would prevent these questions from being asked. Thus, an alternative explanation for the significant reduction in the means with the the multiple-question method is that respondents reported zero to avoid being asked more questions. Table 3.7 shows the percentage of zeros over all items for the before and after-question methods compared to the first question of the multiple-question method. As can be seen, the percentage of zeros reported in the first question of the multiple-question method does not appear to be any higher than the percent in the other two groups. In fact, those answering the first question of the multiple question group reported the lowest percent of zeroes of the three groups. A logit analysis of these data shows no main effect for mode or an interaction effect on the proportion of respondents giving zero responses.

3.3.4 Responses to the Follow-up Questions

Table 3.8 presents an analysis of respondents' answers to the *shoe* and *hours worked* follow-up questions. In the *shoe* question, there was a significant main effect for the order in which the clarifying instructions were presented, but no main effect for mode and no interaction effect between the order of clarification and mode. Contrary to expectations, it appears that respondents were more likely to read and correctly adhere to the clarifying instructions when they were presented after the

Table 3.7: Mean Percentage of Zeros Over All Items (and Sample Sizes) by Method of Presenting Clarifying Instructions and Mode of Data Collection

Mode	Method				Total	χ^2 Statistic
	Clarifying Instructions After	Clarifying Instructions Before	First of Multiple Questions			
	Percent Zeros (n)	Percent Zeros (n)	Percent Zeros (n)	Percent Zeros (n)		
IVR	11.8 (92)	14.3 (95)	8.5 (104)	11.4 (291)	Method	18.29***
Web	10.3 (69)	14.2 (68)	7.2 (69)	10.5 (206)	Mode	< 1 n.s.
Total	11.2 (161)	14.2 (164)	8.0 (172)	11.1 (497)	Interaction	< 1 n.s.

Note: Chi-squared values are from a 3 x 2 logit model and are based on 2 degrees of freedom for method, 1 degree of freedom for mode, and 2 degrees of freedom for the interaction. *** $p < .001$; n.s. denotes not significant.

Table 3.8: Percentage of “Valid” Responses (and Sample Sizes) by Order of Presenting Clarifying Instructions, Mode of Data Collection, and Follow-up Item

Follow-up Item	Mode	Order				χ^2	Statistic	
		Clarifying Instructions After		Clarifying Instructions Before				Total
		Percent Valid	(n)	Percent Valid	(n)			Percent Valid (n)
Shoes	IVR	51.0	(82)	29.0	(86)	39.0 (168)	Order	5.58 *
	Web	47.0	(68)	43.0	(68)	45.0 (136)	Mode	< 1 n.s.
	Total	49.0	(150)	35.0	(154)	42.0 (304)	Interaction	2.59 n.s.
Hours Worked	IVR	80.0	(40)	60.4	(53)	69.2 (91)	Order	< 1 n.s.
	Web	60.8	(51)	75.6	(45)	67.3 (98)	Mode	< 1 n.s.
	Total	68.8	(93)	67.7	(96)	68.3 (189)	Interaction	6.26 *

Note: People who were not employed were excluded from the hours worked item. Chi-squared values for the after vs. before comparison are based on 1 degree of freedom from a 2 x 2 logit analysis. * $p < .05$; n.s. denotes not significant.

shoe question (49%) than when they were presented before (35%).

The *hours worked* item shows a different pattern (Table 3.8). There were no main effects for either method or mode, but there was an interaction effect. Again, contrary to expectations, it seems that respondents were more likely to read and correctly follow the clarifying instructions when they were presented after the *hours worked* question in the IVR (80%) than the Web (61%), but this reversed itself, and respondents were more likely to follow the instructions when they were presented before the question in the Web (76%) than the IVR (60%). With both items, it may

help respondents to follow the instructions when they are the last thing respondents hear before they answer.

3.3.5 Response Times

This analysis is limited to the before/after methods only, since only these methods were comparable in their overall length. As in the previous study, I removed very long times from the analyses.⁵ As hypothesized, respondents tended to spend significantly more time on an item when the clarifying instructions were presented before the question than when they were presented after (see Table 3.9). Eight of the ten items showed this pattern. As can be seen in the last row of Table 3.9, respondents spent nearly 35 seconds more time on average reading and answering the items when the clarifying instructions were presented before rather than after the question.

Four of the ten items also showed significant mode effects and two more of the items showed marginally significant differences. However, the direction of these effects was not as consistent for mode as it was for the method of clarification. Only two of these items (*coats* and *rooms*) moved in the direction predicted. The last

⁵The slowest one percent of values were removed. The cutoff and number of values removed for each item were: residents, 75 seconds (four values); shoes, 92 seconds (four values); coats, 150 seconds (three values); telephone calls, 103 seconds (three values); hours worked, 161 seconds (three values); trips, 97 seconds (three values); emails, 91 seconds (three values); doctor visits, 150 seconds (three values); bedrooms, 65 seconds (three values); rooms, 126 seconds (three values). Removing these times did not change the overall outcome.

Table 3.9: Mean Response Time in Seconds (and Sample Sizes) by Order of Presenting Clarifying Instructions, Mode of Data Collection, and Item

Item	Mode	Order						<i>F</i> -statistic
		Clarifying Instructions After		Clarifying Instructions Before		Total		
		Mean	(n)	Mean	(n)	Mean	(n)	
Resi- dents	IVR	19.9	(93)	31.5	(93)	25.8	(189)	Method 52.01***
	Web	19.0	(67)	22.2	(69)	20.6	(135)	Mode 20.17***
	Total	19.5	(160)	27.6	(164)	23.6	(324)	Interaction 13.95***
Shoes	IVR	25.6	(86)	30.0	(90)	27.8	(176)	Method 9.06**
	Web	27.3	(68)	30.6	(65)	28.9	(133)	Mode < 1 n.s.
	Total	26.3	(154)	30.2	(155)	28.3	(309)	Interaction < 1 n.s.
Coats	IVR	30.2	(81)	40.1	(79)	35.1	(160)	Method 15.78***
	Web	35.6	(67)	43.8	(66)	39.7	(133)	Mode 3.85*
	Total	32.7	(148)	41.8	(145)	37.2	(293)	Interaction < 1 n.s.
Phone Calls	IVR	30.7	(81)	34.1	(77)	32.4	(158)	Method 1.48 n.s.
	Web	33.4	(68)	33.8	(65)	33.6	(133)	Mode < 1 n.s.
	Total	32.0	(149)	34.0	(142)	33.0	(291)	Interaction < 1 n.s.
Hours Worked	IVR	26.4	(80)	32.9	(77)	29.6	(157)	Method 20.46***
	Web	27.3	(68)	35.1	(65)	31.1	(133)	Mode 1.00 n.s.
	Total	26.8	(148)	33.0	(142)	30.3	(290)	Interaction < 1 n.s.
Trips	IVR	27.6	(74)	28.8	(70)	28.2	(144)	Method 4.14*
	Web	27.2	(66)	33.1	(66)	30.2	(132)	Mode 1.32 n.s.
	Total	27.4	(140)	30.9	(136)	28.7	(276)	Interaction 1.90 n.s.
Emails	IVR	24.6	(73)	27.7	(70)	26.1	(143)	Method 1.96 n.s.
	Web	23.4	(66)	24.0	(66)	23.7	(132)	Mode 3.37 †
	Total	24.0	(139)	25.9	(136)	29.1	(275)	Interaction < 1 n.s.
Doctor Visits	IVR	28.0	(71)	32.6	(70)	30.3	(141)	Method 4.57*
	Web	24.7	(67)	27.8	(65)	26.2	(132)	Mode 4.83*
	Total	26.4	(138)	30.3	(135)	24.9	(273)	Interaction < 1 n.s.
Bed- rooms	IVR	19.9	(66)	27.4	(69)	23.7	(135)	Method 29.09***
	Web	16.8	(67)	20.2	(66)	18.5	(133)	Mode 24.83***
	Total	18.4	(133)	23.9	(135)	28.3	(268)	Interaction 3.97*
Rooms	IVR	30.6	(67)	36.8	(69)	33.7	(136)	Method 7.57**
	Web	35.1	(66)	38.9	(66)	37.0	(132)	Mode 3.31 †
	Total	32.8	(133)	37.8	(135)	21.1	(268)	Interaction < 1 n.s.
Sum	IVR	218.0	(93)	254.8	(97)	236.8	(190)	Method 8.07**
	Web	262.1	(69)	294.9	(69)	278.5	(138)	Mode 11.61***
	Total	236.8	(162)	271.5	(166)	254.4	(328)	Interaction < 1 n.s.

Note: Slowest one percent of times for each individual item were removed. All numerator degrees of freedom equal 1 and all denominator degrees of freedom equal the overall n for an item less 4.

row of Table 3.9 gives the average amount of time respondents spent when all ten items were summed. As can be seen, on average, and as predicted, respondents spent nearly 52 seconds more time processing the items when they were presented visually in the Web than when they were presented aurally via IVR.

A couple of the items also showed significant interaction effects (*residents* and *bedrooms*). However, when I summed these over all the items, this effect disappeared (see the bottom panel of Table 3.9).

3.4 Discussion

The results of this study support five main conclusions. The first is that respondents appear to attend to clarifying instructions more when they are placed before the question than when they come afterwards, and they appear to attend to the instructions even more when they are incorporated into a series of questions. Across all ten questions, there was a significant main effect for clarifying method that trended downward (see Panel 1 of Figure 3.1). Previous conversational analyses have suggested that respondents interrupt questions (e.g., Houtkoop-Steenstra, 2002; Oksenberg, Cannell, & Kalton, 1991; Van der Zouwen & Dijkstra, 2002). However, it was not clear whether these interruptions actually led to differences in speakers' answers. A recent finding (reported in Chapter 2) suggests that modest differences in respondents' answers result from placing instructions before a question in a Web survey. The experimental evidence presented in this chapter is congruent with these earlier findings. On average, there was a 5 to 8% reduction in mean responses when

the clarifying instructions were placed before the question relative to when they came after.

The response times also support the conclusion that people pay more attention to instructions that precede the questions. On average, respondents spent 35 seconds less on the ten items when the clarifying instructions followed the questions than when they came before the questions. Thus, respondents' answers to the questions and their response times were congruent with predictions. However, respondents' answers to the follow-up questions were not consistent with expectations. With both follow-up questions, respondents were more likely to correctly exclude the specified subcategories from their responses when the clarifying instructions followed the question than when they preceded it, although this difference was significant only for the *shoe* question. Perhaps this is because both of these follow-up questions applied to individual items for which the method of clarification did not work as expected (see Tables 3.5 and 3.6).

Second, in line with the results presented in Chapter 2, regardless of where the clarifying instructions are presented, they appear to be less effective than asking a series of questions that incorporate the restrictions into the questions. A number of survey researchers have argued that respondents will be more likely to follow instructions if the instructions are incorporated into a series of questions, but there has not been much evidence to support this suggestion (Conrad & Couper, 2004; Conrad & Schober, 2000; Couper, 2008, p. 289; Fowler, 1995, pp. 13-20; Jenkins & Dillman, 1997; Schaeffer & Presser, 2003; Sudman, Bradburn, & Schwarz, 1996, p. 31; Suessbrick, Schober, & Conrad, 2000; Tourangeau, Rips, & Rasinski, 2000,

pp. 38-40, 61). The results of this study offer such evidence. On average, the multiple-question condition resulted in a 15 to 16% reduction in the mean responses compared to the after-question condition. Furthermore, there was some concern that the reduced mean response in the multiple question method might have been the result of deliberate underreporting (that is, respondents may have provided an initial response of zero to keep from being asked additional questions in the multiple question method). However, respondents were no more likely to give initial answers of zero in this method than the other methods (see Table 3.7), suggesting that asking a series of questions does not necessarily encourage underreporting.

A minor issue with the multiple question method was that it elicited more out-of-range values than either of the clarifying methods, probably because it was not evident to respondents that their responses to the sub-questions in the multiple question method were going to yield a negative number. Presumably, when the questions are administered electronically, feedback messages can be used to reduce this problem.

The third main finding was that, in general, the two modes did not affect respondents' understanding of questions with clarifying instructions. This was surprising, given that these instructions were designed to tax respondents' memories in the aural channel. I thought that instructions would be less effective in IVR than the Web, since IVR gives respondents little control over the pace at which the questions are read, and less opportunity to preview and review the clarifying instructions (e.g., Clark & Brennan, 1991; Harwood, 1951; Osada, 2004; Rayner & Clifton, 2009; Schwarz, Strack, Hippler, & Bishop, 1991a). As shown in Panel a of Figure 3.1,

the net result was in the hypothesized direction (higher reports in the IVR than in the Web), but this difference was not significant. In addition, respondents were no more likely to report the correct subcategories in the Web than the IVR (see Table 3.8). Although respondents completed the IVR questions at a significantly faster pace than the Web (see Table 3.9), their answers did not differ by mode. Perhaps the reason for this was that the rate at which the items were delivered in the IVR were still within the limits of normal speech.

Still, differences were found between the modes when it came to the out-of-range values. On average, nearly 7% of the responses were high outliers or negative values in the IVR, whereas a little less than 2.5% of the responses were out of range in the Web condition. IVR respondents may have given more out-of-range responses than Web respondents because they could not see their responses after entering them in IVR, whereas Web respondents could see their responses.

The fourth main finding was the absence of an interaction between the method of clarification and mode. I hypothesized that differences between the modes would be largest when the clarifying instructions followed the questions and smallest when the clarifications were incorporated into the questions. I expected that incorporating the clarifications into a series of questions would equalize the working memory burden in both modes. As shown in Panel a of Figure 3.1, although the results clearly moved in the hypothesized direction, the differences were not significant.

The final main finding of this study is that the number of episodes to be recalled may play a moderating role in determining the effect of the clarifying instructions. The clarifying methods seem to vary in effectiveness in the expected way when the

reported counts are low, but not when the counts are high.

Research on reporting quantities may explain this finding. A number of factors have been shown to affect strategies for answering behavioral frequency questions (e.g., Blair & Burton, 1987; Conrad, Brown, & Cashman, 1998; Menon & Yorkston, 2000). These include the regularity with which the behaviors occur, their similarity to one another, and the frequency with which they occur. When frequencies are low, respondents tend to retrieve and enumerate episodes, but when frequencies are high, they tend to estimate the number of episodes. Perhaps placing the clarifying instructions before the question or asking multiple questions is more effective when the questions elicit low counts because the instructions provide helpful cues that allow respondents to recall infrequent and memorable occurrences once the instructions are brought to their attention. It seems highly plausible, for example, that respondents could enumerate residents and subtract out their children in the residency question, once this clarification was brought to their attention. The fact that respondents answers were altered by the clarifying methods in the low count items is consistent with a conjecture by Schwarz & Oyserman (2001), who speculated that separating a general question into several more specific ones is useful when the specific questions pertain to infrequent and memorable episodes.

Conversely, it may be harder for respondents to implement the clarifying instructions in the high-count items when their answer is based on the overall rate because these questions concern more frequent and less memorable episodes. For example, respondents may not be able to separate out the number of emails they have written for work from those they have written for personal reasons when their

answer is based on the overall rate of sending emails.

Indirect evidence suggesting that respondents may be using different strategies to answer the low versus high-count items comes from an analysis of the proportion of rounded numbers reported in each type. If respondents are having difficulty answering in the high count items and are providing estimates as a result, we would expect to see a larger proportion of rounded values in these items. Conversely, if respondents are enumerating in the low-count items, we would expect to see a smaller proportion of rounded values in these items. The proportion of rounded values in the high count items (.61) is greater than the proportion of rounded values in the low count items (.23).⁶

3.5 Conclusions and Future Studies

Previous mode studies have not shown whether questions with clarifying instructions in mixed-mode surveys should be posed similarly in both modes or tailored to the mode (De Leeuw, 2005; Dillman, 2000; Martin, Hunter Childs, DeMaio, Hill, Reiser, Gerber, Styles, & Dillman, 2007a). At first blush, it may seem as though different perceptual channels will make different demands on comprehension, especially when complex clarifying instructions are involved, so that different methods of providing clarification should be used in the different modes. Consistent with this, the American Community Survey asks a series of residency questions in the interviewer-administered modes of the survey, but a single question with the clarifi-

⁶This difference remains even if the hours worked item is dropped. The most typical responses to this item were zero and 40.

cations (presented after the question) in the self-administered mode. The results of the research presented here, however, suggest that the same methods for clarifying questions should be used in both modes; differences between the channels do not appear to affect respondents' understanding of these instructions. It may be that respondents do not take advantage of the features afforded by the visual mode, such as the ability to more easily review the clarifying instructions. Or maybe the differences between the modes (such as pace) do not move beyond some acceptable range. A fruitful area for future research would be to examine the underlying processes in each channel more closely— for example, how often do respondents return to re-read the clarifying instructions or to listen to them a second time?

The findings reported here suggest that survey practitioners should avoid “tailoring” clarifying methods in mixed-mode surveys in suboptimal ways. Otherwise, differences in response means may be obtained between the different modes of a mixed-mode survey that are not attributable to differences in channel, but due instead to employing a sub-optimal clarifying method in one of the modes. In general, it appears that instructions that are placed before a question in either channel have a somewhat better chance of being attended to than instructions placed after the question; instructions incorporated into a series of questions are even more effective. Still, respondents' answers to the follow-up questions suggested that at least with two of the items that respondents' answered more accurately when the clarifying instructions came after the question than before.

Finally, the key findings appear to depend on the counts reported. When the counts being queried are low, there is some evidence that respondents are recalling

and counting specific episodes, and the manner in which the clarifying instructions are presented can either help or hinder this process. But when the number of articles asked about are high, it appears that respondents may be estimating and the manner in which the clarifying instructions are presented may not matter so much. Further research is needed, however, to explicate the relationship between the method of presenting clarifying instructions, the mode in which they are presented, and the characteristics of the quantities being asked about.

Chapter 4

Varying Examples and Their Presentation in a Web Survey

4.1 Introduction

Survey questions often use examples, that is, subcategories that are meant to clarify the intent of a question's concepts. Previous research suggests that examples have a positive effect (an increase in the level or detail in reporting) (e.g., Martin, 2002; Martin, Sheppard, Bentley, & Bennett, 2007b; Tourangeau, Conrad, Couper, Redline, & Ye, 2009; Tourangeau, Conrad, Couper, & Ye, 2010). For instance, Tourangeau, Conrad, Couper, & Ye (2010) found that informing respondents that grain includes such examples as *bread*, *pasta*, and *rice* led to a higher reported consumption of grain products. Presumably, the examples cue respondents to include instances they would not otherwise have reported. The finding that respondents' report higher frequencies when given examples compared to their absence is consistent with the "unpacking effect" (Tversky & Koehler, 1994). According to Tversky & Koehler (1994), an unpacked description of an event (e.g., a businessman does business with England, France, or some other European country) elicits higher probability judgments than a packed description (e.g., a businessman does business with a European country).

Not only does the presence of examples appear to have an effect, but their characteristics appear to play a role as well. Sloman, Rottenstreich, Wisniewski,

Hadjichristidis, & Fox (2004) proposed that the typicality of the examples presented is a key determinant in the unpacking effect. The researchers compared typical and atypical examples and showed that the increase in rated probability found in the unpacking studies disappeared when the category was unpacked into typical examples only. Sloman and his colleagues (2004) argue that people judge categories in terms of their typical instances to begin with so that providing typical examples does little to affect their judgments. Furthermore, the researchers found that atypical examples *decreased* probability judgments compared to typical examples or no examples at all. They argue that atypical examples inhibit people from thinking of and including more typical examples. This inhibiting effect is known as part-set cueing in the memory literature (e.g, Roediger, 1974).

Tourangeau, Conrad, Couper, & Ye (2010) also compared typical (e.g., *lettuce, tomatoes, carrots*) and atypical (*French fries, potato chips, onions*) examples in survey questions; in general, their results are in line with those found in the unpacking studies of probability judgments. These investigators found an effect for the presentation of atypical examples in survey questions. For instance, respondents reported consuming 9.7 servings of vegetables when they got atypical examples compared to 8.2 servings in the no example group and 8.9 servings in the typical examples group, suggesting that respondents may expand their interpretations of a category when they are given atypical examples. Tourangeau, Conrad, Couper, & Ye (2010) found an effect for another characteristic of the examples as well— the frequency with which the examples were consumed. Respondents who got frequently consumed examples of vegetables (e.g., *lettuce, tomatoes, and carrots*) consistently reported

greater consumption of vegetables than those who got infrequently consumed examples (e.g., *asparagus*, *Brussel sprouts*, and *green beans*). Respondents seem to focus on high frequency foods when they are given high frequency examples and on low frequency foods when they are given low frequency examples.

An unexplored variable to date is the breadth of the category asked about. In Tourangeau, Conrad, Couper, & Ye (2010), the categories under study varied in their breadth. For example, their study included dairy and grain products, which are broad categories (in a hierarchical ordering of food products, dairy and grain products would top the list). They also included poultry, which represents a narrower category (poultry is a subcategory under meat). The research question is whether the breadth of the category interacts with the frequency of the examples, such that broad categories and frequent examples elicit disproportionately more reports than narrow categories and infrequent examples.

Another unexplored area is whether it matters where and how the categories and examples are presented. Previous research with clarifying instructions showed that presenting such instructions before a question increases their effectiveness somewhat relative to presenting them after the question (see Chapters 2 and 3). I hypothesized that a similar effect would occur when the order in which a short list of examples is varied. In previous research, the effect of clarification method appeared to interact with the overall frequency reported. Placing the clarification before the question appeared more effective than placing it after, but this effect seemed to disappear when the frequencies to be reported were high (see Chapter 3). I hypothesized that a similar interaction might be seen between the frequency of examples

and order in this study.

Manipulating the font of the lengthy clarifying instructions did not affect respondents' interpretations previously (see Chapter 2), and so I expected that the same null effect would carry over to examples. However, some research finds that respondents pay more attention to instructions when they are placed in bulleted format than in text format (e.g., Hartley, 2004). Therefore, I hypothesized that a bulleted format might increase the impact of the examples and lead to higher reports of consumption; I also thought this effect might be enhanced even further when the bulleted examples came before the question than after.

This chapter describes the results of an experiment that was designed to study five variables that might influence a respondents' interpretation of food categories in survey questions and their reported consumption of food in these categories. The high frequency examples were expected to increase reported consumption, especially with the broad food categories. The other three variables were expected to affect the impact of the examples. Two of these (order and format) were expected to increase the effect of the examples (presenting examples before the question and in bulleted format was expected to increase the level of reporting). Based on earlier studies, I expected the font of the examples to have little impact.

4.2 Methods

4.2.1 Study Questions and Experimental Conditions

This experiment examined four food frequency items in a Web survey. These items asked respondents how many servings they had eaten from one of four major food groups (meat, dairy, grain, and fruits and vegetables). The basic question about meat consumption took this form: “For the purposes of this question, meat includes beef, pork, poultry, and other meat. How many servings of meat do you typically eat each week?”

The experiment varied five factors. The first factor was the breadth of the food category. Some participants answered questions about a broad category (such as meat). Others answered questions about narrower categories (such as poultry). The second factor varied the frequency with which the examples presented were eaten (frequently versus infrequently). The frequency with which the examples are consumed has been shown to affect the reported frequency of the overall category (Tourangeau, Conrad, Couper, & Ye, 2010). The question on meat consumption above mentions frequently eaten examples *beef*, *pork*, and *poultry*, whereas examples of infrequently eaten meat would include *lamb*, *veal*, and *goat*.

Table 4.1 shows how each of the food categories (broad versus narrow) and examples within categories (high versus low frequency) were varied for each of the individual items. A national food coding scheme was used to categorize food categories as broad or narrow (Bodner-Montville, Ahuja, Ingerwersen, Haggery, Enns, & Perloff, 2006), and a national nutritional survey was used to classify each example

Table 4.1: Examples by Breadth and Frequency of Consumption

Breadth	Categories	Frequency	
		Frequent Examples	Infrequent Examples
Broad	Meat	Beef, pork, poultry	Lamb, veal, goat
Narrow	Poultry	Chicken, turkey, duck	Goose, quail, pheasant
Broad	Dairy Products	Milk, cheese, yogurt	Frozen yogurt, feta cheese, custard
Narrow	Cheese	Cheddar cheese, Swiss cheese, cottage cheese	Blue cheese, Brie, Gouda
Broad	Grain Products	Bread, pasta, rice	Millet, puffed wheat, couscous
Narrow	Bread	White bread, French bread, corn bread	Wheat bread, whole grain, brioche
Broad	Fruits & Vegetables	Apples, bananas, lettuce	Grapefruit, dried fruit, asparagus
Narrow	Vegetables	Lettuce, tomatoes, carrots	Asparagus, Brussel sprouts, green beans

as frequently or infrequently consumed.¹

The remaining factors varied the manner in which the categories and examples were presented. The first of these, shown in Figure 4.1, was the order of presentation (presenting the examples either before or after the question). The second was the font in which the examples were presented (either presented in the same font as the questions or in italics). The final presentation factor was the format in which the examples were presented (presenting them vertically in bulleted format or as text).

Overall, the design formed a 2x2x2x2x2 factorial experiment with a total of

¹The National Health and Nutrition Examination Survey 2003-2004 Dietary Interview

A. Before, Same Font, Bullets

For the purposes of this question, meat includes:

- beef,
- pork,
- poultry,
- other meat.

How many servings of meat do you typically eat each week?

servings

B. After, Same Font, Bullets

How many servings of meat do you typically eat each week?

For the purposes of this question, meat includes:

- beef,
- pork,
- poultry,
- other meat.

servings

C. Before, Italics, Bullets

For the purposes of this question, meat includes:

- *beef,*
- *pork,*
- *poultry,*
- *other meat.*

How many servings of meat do you typically eat each week?

servings

D. After, Italics, Bullets

How many servings of meat do you typically eat each week?

For the purposes of this question, meat includes:

- *beef,*
- *pork,*
- *poultry,*
- *other meat.*

servings

E. Before, Same Font, Text

For the purposes of this question, meat includes beef, pork, poultry, and other meat. How many servings of meat do you typically eat each week?

servings

F. After, Same Font, Text

How many servings of meat do you typically eat each week?

For the purposes of this question, meat includes beef, pork, poultry, and other meat.

servings

G. Before, Italics, Text

For the purposes of this question, meat includes beef, pork, poultry, and other meat. How many servings of meat do you typically eat each week?

servings

H. After, Italics, Text

How many servings of meat do you typically eat each week?

For the purposes of this question, meat includes beef, pork, poultry, and other meat.

servings

Figure 4.1: The Order, Font, and Format of the Examples for One Item

32 cells (see Appendix A.3 for the four major questions expressed in all 32 ways).

4.2.2 Data Collection and Sample

This experiment was embedded in a Web survey that was conducted by Market Strategies International (MSI) from June 23, 2010 through July 21, 2010. One half of the sample for this survey was drawn from the Survey Sampling International's (SSI) online panel. This panel is an opt-in panel of over 1.3 million persons who have signed up to receive survey invitations. The other half of the sample was drawn from the three-million member Authentic Response panel.

Respondents from both panels were invited by email to participate in the study. Each respondent was assigned a unique PIN that allowed access to the survey. SSI panelists were offered a sweepstakes incentive to take part and Authentic Response panelists were offered an incentive of \$0.75. Non-respondents received one follow-up email in both samples.

A total of 2,407 respondents completed the survey. SSI selected 138,323 of its panel members and 1,201 completed the survey, for a 1.0% response rate. Authentic Response contacted 15,435 panel members, of whom 1,206 completed the survey, for an 8.0% response rate. The overall response rate was 2.0% (AAPOR RR1).

Respondents were randomly assigned to one of the 32 conditions, and once assigned to a condition, all four questions were asked in the style of that condition. Altogether, the survey included 75 questions. The four questions from this particular experiment were questions 23 through 26 in the questionnaire.

4.3 Results

The analysis examines two outcome variables— respondents’ answers to the questions (the mean number of servings consumed in a typical week) and the time it took respondents to answer the questions (in seconds). Respondents’ answers to the study questions included a few very high values. For example, one respondent reported consuming 555 servings of meat in a week. Values greater than 50 were removed; an inspection of the items’ distributions suggested that these values were outliers reflecting respondent errors.²

4.3.1 Mean Number of Servings

Table 4.2 shows the mean number of servings consumed in a typical week by item and by experimental condition. As expected, the breadth of the target category had a large effect on respondents’ answers. Across all four items, respondents consumed an average of 9.5 servings of food per week when the categories were broad compared to 5.9 servings when the categories were narrow.

The frequency with which the examples are usually consumed also had an effect on the reports. Respondents consistently reported consuming more servings of food when the examples were frequently consumed foods than when they were infrequently consumed foods. For instance, respondents reported consuming more servings of meat when they received *beef*, *pork*, and *poultry* as examples (5.5 servings)

²Three observations greater than 50 were removed for the meat item, four for dairy, seven for the grain item, and three for the vegetable/fruit item. Removing these outliers did not alter any of the conclusions.

Table 4.2: Mean Number of Servings (and Sample Sizes) by Item and Experimental Condition

Condition	Number of Servings				
	Meat/ Poultry	Dairy/ Cheese	Grain/ Bread	Vegetables& Fruits/ Vegetables	Average
	Mean (n)	Mean (n)	Mean (n)	Mean (n)	Mean (n)
Breadth					
Broad	7.04 (1213)	9.52 (1210)	9.46 (1207)	11.61 (1210)	9.42 (1216)
Narrow	3.60 (1186)	4.50 (1184)	6.78 (1186)	8.69 (1185)	5.90 (1187)
Frequency					
Frequent	5.61 (1231)	7.29 (1229)	8.40 (1231)	10.53 (1229)	7.97 (1233)
Infrequent	5.05 (1168)	6.76 (1165)	7.84 (1162)	9.78 (1166)	7.37 (1170)
Order					
Before	5.41 (1211)	7.03 (1209)	8.24 (1210)	10.13 (1211)	7.72 (1233)
After	5.27 (1188)	7.04 (1185)	8.02 (1183)	10.20 (1184)	7.64 (1170)
Format					
Bullets	5.36 (1196)	7.06 (1190)	8.19 (1190)	10.07 (1193)	7.67 (1196)
Text	5.32 (1203)	7.00 (1204)	8.08 (1203)	10.26 (1202)	7.68 (1207)
Font					
Same	5.41 (1166)	7.09 (1163)	8.08 (1163)	10.11 (1165)	7.68 (1168)
Italics	5.27 (1233)	6.99 (1231)	8.18 (1230)	10.22 (1230)	7.68 (1235)
F-statistic					
Breadth	383.64 ***	485.33 ***	101.74 ***	79.51 ***	365.05 ***
Frequency	9.97 **	5.38 *	4.42 *	5.31 *	10.51 **
Order	0.05 n.s.	0.45 n.s.	0.32 n.s.	0.26 n.s.	0.20 n.s.
Format	0.06 n.s.	2.84 n.s.	0.09 n.s.	0.55 n.s.	0.19 n.s.
Font	0.58 n.s.	0.23 n.s.	0.20 n.s.	0.09 n.s.	0.00 n.s.

Note: Values greater than 50 were removed. For the F -statistics, the numerator degrees of freedom are 1 and the denominator degrees of freedom are the sample sizes for the item (approximately 2400) less 32. *** $p < .001$, ** $p < .01$, * $p < .05$, n.s. denotes not significant.

than they did when they received *lamb*, *veal*, and *goat* as examples (5.1 servings). Again, this same pattern was evident for all four items. Across all four food categories, respondents reported consuming an average of almost 8.0 servings of food when they got frequently consumed foods as examples; this dropped to an average of 7.4 servings when they got infrequently consumed foods as examples.

Contrary to expectations, neither order nor format had a main effect on respondents' answers. Finally, whether the examples were in the same font as the questions or italics had no main effect either.³

4.3.2 Response Times

As in the earlier studies, I removed the slowest one percent of the times for each of the items in this analysis. Twenty four values were removed from the meat item, 23 from the dairy item, 26 from the grain item, and 25 from the vegetables/fruits item. Table 4.3 shows the mean response time by item and the five main conditions. Respondents consistently took longer to answer a question when the question asked about a broad category than a narrow one. This difference was significant for all but one of the items, the vegetables/fruits category. On average, respondents took nearly five seconds more to answer all four of the questions when the categories were broad than when they were narrow (54.7 seconds versus 49.0 seconds); this result was highly significant overall and for three of the four items.

Respondents also consistently took longer to answer a question when it in-

³Across all questions, a few three-way interactions emerged over all questions, but they were scattered and did not appear to have any meaningful interpretations.

Table 4.3: Mean Response Time in Seconds (and Sample Sizes) by Item and Experimental Condition

Condition	Response Time				
	Meat/ Poultry	Dairy/ Cheese	Grain/ Bread	Vegetables& Fruits/ Vegetables	Total
	Mean (n)	Mean (n)	Mean (n)	Mean (n)	Mean (n)
Breadth					
Broad	13.1 (1206)	13.2 (1206)	14.7 (1207)	14.2 (1204)	54.7 (1219)
Narrow	11.5 (1176)	11.2 (1177)	13.1 (1173)	13.7 (1177)	49.0 (1188)
Frequency					
Frequent	11.7 (1226)	12.0 (1216)	13.4 (1221)	13.5 (1222)	50.1 (1235)
Infrequent	12.9 (1156)	12.5 (1167)	14.5 (1159)	14.5 (1159)	53.8 (1172)
Order					
Before	12.5 (1203)	12.4 (1204)	14.1 (1208)	13.7 (1203)	52.3 (1217)
After	12.1 (1179)	12.1 (1179)	13.7 (1172)	14.2 (1178)	51.5 (1190)
Format					
Bullets	12.3 (1187)	12.3 (1183)	13.8 (1182)	13.9 (1198)	51.9 (1197)
Text	12.3 (1195)	14.1 (1200)	14.1 (1198)	14.0 (1183)	51.9 (1210)
Font					
Same	12.7 (1158)	12.4 (1161)	14.0 (1158)	14.3 (1157)	52.8 (1171)
Italics	12.0 (1124)	12.0 (1222)	13.9 (1222)	13.7 (1224)	51.0 (1236)
<i>F</i>-statistic					
Breadth	23.31 ***	30.75 ***	11.53 ***	1.46 n.s.	25.0 ***
Frequency	12.51 ***	1.84 *	4.88 *	4.49 *	10.6 ***
Order	1.68 n.s.	0.77 n.s.	0.72 n.s.	0.95 n.s.	0.4 n.s.
Format	0.08 n.s.	0.01 n.s.	0.38 n.s.	0.09 n.s.	0.1 n.s.
Font	4.66 *	1.11 n.s.	0.01 n.s.	1.63 n.s.	2.3 n.s.

Note: The slowest one percent of times were removed. For the *F*-statistics, the numerator degrees of freedom are 1 and the denominator degrees of freedom are the sample sizes for an item less 32. *** $p < .001$, ** $p < .01$, * $p < .05$, n.s. denotes not significant.

cluded infrequently consumed examples than frequently consumed ones. This difference was significant for each of the items individually and overall. Overall, respondents took nearly four seconds longer when the examples were infrequently consumed (mean of 53.8 seconds) than when they were frequently consumed (mean of 50.1 seconds).

As predicted, respondents appeared to spend more time on the question when the examples were before the question text. They also appeared to spend more time when the examples were in the same font as the questions; however, neither main effect was this significant, either overall or for any of the individual items. The one exception occurred for the font of the meat item. Respondents did spend significantly more time on that item when the examples were in the same font as the question than when they were italicized.

The amount of time a respondent spent when the examples were presented in a bullet format did not conform with expectations. For three of the four items (*dairy*, *grain*, and *vegetables and fruits*), the time respondents spent on an item was in the opposite of the expected direction – less time spent on the bulleted examples than when the examples were provided as ordinary text. However, these differences were not significant. For the meat item and over all items, the means were identical.

An interaction emerged between the order and format variables ($F(1, 2375) = 5.82, p < .05$). Placing the examples before the question in text format appeared to increase the amount of time respondents spent on the question (53.7 seconds for the before/text group versus around 50.0 seconds for the other three groups).

4.4 Discussion

The clearest finding from this experiment is that people's answers depend on the frequency of the examples they are given, but not on how the categories or the examples are presented. In addition, respondents report much larger mean frequencies and take longer to answer when a question asks about a broad category (e.g., meat) than about a narrower one (e.g., poultry). Respondents doubtless consider a larger set of instances when the category is broad, and spend more time doing so than when the category is narrow.

As expected, respondents appear to report larger mean frequencies when they get frequently consumed examples rather than infrequently consumed ones. This finding replicates that of Tourangeau, Conrad, Couper, & Ye (2010) (see also Tourangeau, Conrad, Couper, Redline, & Ye, 2009). However, Tourangeau, Conrad, Couper, & Ye (2010) also found that respondents took a few seconds more to report when they got frequent examples (53.9 seconds across all four questions in their study) than when they got infrequent ones (52.0 seconds). In this study, the reverse was true, even though the amount respondents reported was larger in both studies. One explanation for this reversal may be that Tourangeau, Conrad, Couper, & Ye (2010) crossed the frequency with which examples were consumed with their typicality, and one of these combinations (frequent-atypical examples) appeared to slow people down. For example, providing *French fries*, *potato chips*, and *onions* as examples of vegetables appeared to increase response times. In my study, respondents may not have been similarly slowed because, in general, the examples were

typical. For instance, both the frequent examples of vegetables (*lettuce, tomatoes, and carrots*) and infrequent examples of vegetables (*asparagus, Brussel sprouts, and green beans*) were typical in my study. Thus, in my study, frequently consumed (and typical) examples may have been more accessible – that is, easier and quicker for people to recall – than infrequently consumed examples.

The presentation variables had little effect on respondents' answers or the time it took them to formulate them. Given the results of previous research (see Chapter 1), I didn't expect that changing the font would have an effect, but I did think that placing the examples before the question (see Chapter 2) and presenting them as bullets might facilitate processing of this information. Overall, order did work in the predicted direction (both in terms of respondents' answers and response times), but the differences were not significant. The results of the format variable (bulleted versus text) were both small and inconsistent, and were not significant either. Order and format did interact to affect response times. Presenting the examples before a question in text format enhanced the amount of time respondents spent on an item, but this did not have an effect on the number of servings they reported.

The implication of this research for survey practitioners is that the choice of examples requires careful consideration, as this choice will influence respondents' answers, but it appears survey practitioners need worry little about where and how to present examples.

Chapter 5

Conclusions

5.1 Summary

Many concepts are inherently ambiguous or vague. I began this dissertation with a striking illustration from recent events to emphasize the point that even concepts that appear to be straightforward initially – for example, the concept of the “the 9/11 Twin Towers” – can turn out to be ambiguous. It may not seem so at first, but the attack on the World Trade Center can be viewed as one event or two, depending on one’s perspective. Chapter 1 shows that this same phenomenon extends to survey questions. For example, in a study that examined a Current Population Survey supplement question, Suessbrick, Schober, & Conrad (2000) showed that “smoking cigarettes” meant “only cigarettes you finished,” to some respondents, but “even just one puff” to others. Survey questions often provide definitions, instructions, or examples to reduce the ambiguity or vagueness of survey concepts (e.g., Conrad, Schober, & Coiner, 2007; Tourangeau, Conrad, Couper, & Ye, 2010). However, presenting such information effectively poses a challenge of its own, as respondents do not always attend to this clarification and when they do, it may not have the desired effect (e.g., Tourangeau, Conrad, Arens, Fricker, Lee, & Smith, 2006). The main aims of my dissertation were to examine whether respondents attend to clarifying instructions and examples in survey questions and to explore

conditions that might increase the effectiveness of such material in clarifying survey questions. I conducted three experiments to identify methods for improving respondents' understanding of concepts in survey questions. The first experiment examined the effect of providing clarifying instructions in a Web survey. The second experiment extended this comparison to include a second mode of data collection, the IVR, and thus compared aural and visual presentation of clarifying information. And the third experiment studied the effect of providing examples in a Web survey. The results from these experiments are reported in Chapters 2 through 4 of this dissertation.

Experiment 1 showed that answers changed in the expected direction when clarifying instructions accompanied a question. Eight questions, patterned after questions in federal surveys, were administered. The questions focused on the household (e.g., the number of residents) or items owned by the respondent (e.g., the number of shoes). The instructions that accompanied these questions directed respondents to exclude particular subcategories. For example, a question about shoes was posed: *How many pairs of shoes do you own? For the purposes of this question, do not include boots, sneakers, athletic shoes, or bedroom slippers. Include sandals, other casual shoes, and dress shoes. If you do not own a pair of shoes (as we have defined them), enter "0."*

The clarifying instructions (whether they came before or after a question) reduced the mean response by about 20%. However, these instructions were not completely successful. Respondent answers to the follow-up *shoe* question, which was designed to determine whether respondents were excluding the correct subcategories

from their responses, indicated that respondents made the correct exclusions only about 50% of the time. Furthermore, clarifying instructions were only about half as effective as asking multiple questions. Asking multiple questions reduced the mean response by about 33% compared to the questions without instructions. Assuming that lower responses are more accurate, the multiple question method is more effective at eliciting accurate responses than providing clarifying instructions. Numerous researchers have speculated as much, but little empirical research existed to confirm this speculation (Conrad & Couper, 2004; Conrad & Schober, 2000; Couper, 2008, p. 289; Fowler, 1995, pp.13-20; Jenkins & Dillman, 1997; Schaeffer & Presser, 2003; Sudman, Bradburn, & Schwarz, 1996, p. 31; Suessbrick, Schober, & Conrad, 2000; Tourangeau, Rips, & Rasinski, 2000, pp. 38-40, 61).

Both Experiments 1 and 2 indicate that some methods of providing clarifying instructions are more effective than others. One such variable is whether the clarifying information is given to respondents before or after a question. Respondents seem to anticipate the end of a question and they often prepare to answer before they hear the end of the question or anything that comes afterward (e.g., Houtkoop-Steenstra, 2002; Oksenberg, Cannell, & Kalton, 1991; Van der Zouwen & Dijkstra, 2002). Respondents may interrupt questions, but it is not clear whether this has an impact on their answers. Again, assuming that lower answers are more accurate, the results of Experiments 1 and 2 suggest that respondents pay somewhat more attention to instructions that are presented before a question than after (see Tables 2.3 and 3.5). For example, presenting the clarifying instructions before the question resulted in an 5% reduction in respondents' answers over all questions in

Experiment 1 and a 4% reduction in Experiment 2 compared to presenting them after the question. In addition, respondents spent about 3 to 6 seconds more per question when the instructions came before the question than after it (see Tables 2.5 and 3.9).

In general, the main results from Experiment 2 were consistent with those in Experiment 1. Presenting the clarifications after the question produced the highest mean response, presenting the instructions before produced an intermediate mean, and asking a series of questions produced the lowest mean. Respondents' answers to the questions and their response times in those two studies seemed congruent with each other and with the predictions. However, the results from the follow-up questions that probed respondents' answers were more mixed. I thought that respondents would report fewer of the excluded subcategories when the clarifying instructions were presented before the question than after. However, there was only one significant difference between the before and after conditions in the percent reporting valid subcategories out of four comparisons, and that difference was in the opposite from the predicted direction. The inconsistency between the follow-up analysis and the other analyses may be attributable to the small sample sizes for the follow-up probes. Or it may be due to the fact that the clarification methods did not work as expected in the first place for the *shoe* and *hours worked* questions.

Experiments 1 and 2 also showed that neither the font in which the clarifying instructions appeared nor the mode of data collection significantly affected respondents' answers. I hypothesized that respondents would pay more attention to the clarifying instructions if they were in the same font as the question and tend to

ignore them when they were italicized, but it appears that both their answers and the time it took them to formulate an answer was about the same whether the clarifications were italicized or not.

The aural modes of data collection (such as IVR) make it harder for respondents to control the pace of question delivery and to review the questions. These are often cited as reasons to keep the response task simple in those modes. If we think of the clarifying instructions as part of the question, the questions can be viewed as long and the tasks complex, requiring numerous calculations and making onerous demands upon memory. However, the answers of the IVR respondents did not differ significantly from those of the Web respondents over all questions in Experiment 2, suggesting that there are no overall differences in cognitive demand (or burden) between the two channels. Both readers and listeners benefit from questions that are broken down into multiple, simpler components. It also appears that both readers and listeners benefit modestly from having the clarifying methods provided to them before the question rather than after.

Further analyses of the results of Experiment 2 suggested that these findings depended on the underlying frequency to be reported. In high count questions (whose means were above 8, such as the number of phone calls made), respondents appeared to estimate more (e.g., they were more likely to provide a rounded response) and their answers did not vary with the clarification method. However, for reasons that are unclear, respondents also provided lower estimates in the IVR for the high count items than they did in the Web.

The final experiment showed that for the most part, respondents paid attention

to a short list of examples, regardless of where or how they were presented. An example of a question with examples is: *How many servings of meat do you typically eat each week? For the purposes of this question, meat includes beef, pork, poultry, and other meat.* The findings here – that none of the presentation variables under study (order, font, or format) were significant – seems consistent with those reported in earlier chapters. Although the order in which the clarifying information was given was shown to have an effect in Experiments 1 and 2, the clarifications in those studies were more elaborate, yet the effect was still relatively small for both of those studies. Consistent with this, the less elaborate clarifications of Experiment 3 led to no effect.

The example experiment was the only one to vary the format of the clarifying information. Although I hypothesized that respondents would attend to the examples more when they were presented in a bulleted format than when they were presented as text, it appears that presenting such a short list of examples in the form of bullets did not make them any more effective.

A final finding to come out of Experiment 3 was that while the presentation variables did not affect respondents' answers, the content variables did. As predicted, broad categories elicited higher reports than narrow categories and frequently consumed examples elicited higher reports than infrequently consumed examples. Taken together, the findings suggest that survey practitioners need be concerned with the content of the examples more than their presentation.

Table 5.1 summarizes the conditions that were common across all three experiments. It illustrates the impact of adopting the before or multiple question methods of clarification in all three experiments compared to using the after method. In this

table, the means of the after method are set to one, and the other entries in the table represent the mean over all questions in that method compared to means across all questions in the after condition. Respondents' answers moved in the expected directions: in the first two experiments, respondents' answers in the before and multiple questions conditions are a fraction of what they were in the after condition. In the third experiment, they were slightly higher in the before condition than the after condition. Table 5.1 also shows that the impact of the before condition is small in comparison to the multiple question condition. The before condition ranges from .05 less to .01 larger than the after condition. But the multiple question condition ranges from 0.22 to 0.31 less than the after condition. Thus, when getting people to pay attention to the clarifying instructions, no method of presenting instructions compares to asking multiple questions, and this applies to both channels of communication. By implication, the practice of asking multiple questions in the aural channel of a survey, such as the American Community Survey, but a single question in the visual channel is not the most effective approach overall. The results of this research suggest that lengthy clarifying instructions should be incorporated, and a series of questions asked in both channels of a mixed-mode survey.

There is no evidence from either of the first two experiments that respondents satisficed (i.e., reported zeros in response to the first question in the multiple question series to avoid being asked more questions). Nor is there any evidence that they went back and changed their answers to the first question in the series after having answered subsequent questions. Finally, there is no evidence that they were more likely to breakoff when asked multiple questions rather than one question with clar-

Table 5.1: Means of the Before and Multiple Question Methods Relative to the After Method for Experiments 1, 2, and 3

Method	Experiment 1	Experiment 2		Experiment 3
	Web	IVR	Web	Web
After	1.00	1.00	1.00	1.00
Before	0.95	0.99	0.95	1.01
Multiple Questions	0.69	0.88	0.86	NA

ifying instructions. Still, there are a few potential downsides to using the multiple question method. For one, it takes longer, on average, to answer a series of questions than a question with clarifying instructions. And chances are, if left unchecked, the multiple question method will generate a small percentage of negative values. In computerized surveys, these negative values can probably be minimized through the use of feedback mechanisms. Ultimately, the multiple question method appears to be the most promising method of improving respondents' understanding of complex concepts, despite these shortcomings.

5.2 Limitations

These experiments have their limitations, especially when looked at in isolation. However, when the experiments are grouped together, some of these limitations seem less troubling. For example, all three experiments generalize to different popu-

lations. The sample in Experiment 1 was designed to represent the U.S. population. Experiment 2 was designed to generalize to the U.S. population of Internet users. And the third experiment was designed to generalize to the opt-in Web respondents who chose to participate. Not all selected persons responded in the first two studies, which threatens their generalizability to their respective populations. However, consistent findings among the three studies should mitigate the concerns regarding the limitations of any one of the studies; the key findings appear to be robust despite sample differences.

Secondly, I restricted myself to the study of factual non-sensitive questions. In the first two studies, the questions concerned the respondent's household or belongings, and in the third study, the questions concerned foods they had recently eaten. It is possible that the findings from these questions will not generalize to other kinds of questions (for example, attitudinal or sensitive questions), although it is not obvious why that would be the case.

A third limitation relates to the way the clarifying instructions were written in the first two experiments. To allow me to predict the direction of the effects of the clarifying instructions, the instructions were written in a way that they may have been seen as surprising or unusual. It may be that the findings do not generalize beyond surprising instructions.

Fourthly, all three studies use the overall means to measure the effectiveness of the instructions. The first two studies assume that a lower answer indicates greater accuracy and the third study assumes that a higher answer indicates greater accuracy. However, in the absence of true scores, we can never be fully confident

that respondents' answers are genuinely more accurate.

Finally, all three experiments were undertaken in computerized self administered modes of data collection. It is unclear how well the findings will generalize to other modes of data collection, especially mail.

5.3 Future Research

Future research should be conducted with additional questions to assess the generalizability of the findings from this dissertation. Also, it would be good to confirm the accuracy of the answers. This could be accomplished by providing respondents with scenarios with which to answer these same items, or asking a different set of items that can be validated through the use of external data. In future research, following up more than the *shoe* and *hours worked* questions would increase the power of this analysis.

A comparison of the clarifying instruction experiments (Experiments 1 and 2) with the example experiment (Experiment 3) suggests that the length of the clarifying information may matter. Future research should examine this possibility.

Left for future research is a more direct assessment of how respondents process clarifying instructions in the two sensory channels. Partly, this could be done with an eye-movement analysis. Do respondents skip over the clarifications in Web surveys, but return to them when needed? How do these eye-tracking measures compare with respondents' interrupting the reading of clarifications in the aural channel and asking for questions to be repeated? It may be interesting to compare questions

with clarifications that are intended to apply to everyone (such as those studied in this dissertation) with those that are expected to apply to only a few respondents. Do respondents have a greater tendency to ignore the clarifications in the latter case?

Also, left for future research is how well the Web survey findings here translate to mail surveys, especially when it comes to the multiple question method. Asking multiple questions in the Web is easier than in paper, since Web surveys can be programmed to automatically navigate respondents through complex skip patterns. It is also possible that unit nonresponse rates will be even larger in mail surveys than either Web or IVR, since respondents can see that more questions are being asked.

Finally, a useful follow-up to the experiments I conducted here would be to compare respondents' answers from differing modes of a mixed mode survey, such as the mail mode of the American Community Survey with those from CATI, after taking into account respondents' self-selection into these modes and the differences in clarification methods used by the two modes. These analyses look at more than channel effects, since CATI involves interviewer administration, but it would nonetheless be useful to see if differences that are apparent in an ongoing federal survey are similar to those found in these experiments. These issues, and more, are left for future research.

Appendix A

Questionnaires

A.1 Experiment 1

[Experiment A: Questions with Lengthy Clarifying Information

This is a one-factor experiment. Present one item per screen with labeled numeric entry box. There are 10 items in this experiment, blocks of which are interspersed between experiments 1 through 4. The ten items are: 1, 8, 9, 10, 12, 13, 17, 18, 26, and 27. Once a respondent is randomly assigned to one of the six conditions below, the respondent should receive all ten items in that one condition.

Experiment A

1. One Question, No Qualifications
 2. One Question, Qualifications After, Same Font
 3. One Question, Qualifications Before, Same Font
 4. One Question, Qualifications After, Italics
 5. One Question, Qualifications Before, Italics
 6. Multiple Questions with Qualifications Incorporated]
-

[IF EXPERIMENT A = 1 ASK 1A]

1A .The first question is about the people at this address.

How many people are currently living or staying at this address?

Number of people

[IF EXPERIMENT A = 2 ASK 1B]

1B. The first question is about the people at this address.

How many people are currently living or staying at this address? Do not forget to count yourself. For the purposes of this question, a person is defined as someone 18 years or older. Do not include children 17 years or younger. Do not include anyone who is living somewhere else for more than 2 months, such as a college student living away or someone in the Armed Forces on deployment.

Number of people

[IF EXPERIMENT A = 3 ASK 1C]

1C. The first question is about the people at this address.

Do not forget to count yourself. For the purposes of this question, a person is defined as someone 18 years or older. Do not include children 17 years or younger. Do not include anyone who is living somewhere else for more than 2 months, such as a college student living

away or someone in the Armed Forces on deployment. How many people are currently living or staying at this address?

Number of people

[IF EXPERIMENT A = 4 ASK 1D]

1D. The first question is about the people at this address.

How many people are currently living or staying at this address? *Do not forget to count yourself. For the purposes of this question, a person is defined as someone 18 years or older. Do not include children 17 years or younger. Do not include anyone who is living somewhere else for more than 2 months, such as a college student living away or someone in the Armed Forces on deployment.*

Number of people

[IF EXPERIMENT A = 5 ASK 1E]

1E. The first question is about the people at this address.

Do not do not forget to count yourself. For the purposes of this question, a person is defined as someone 18 years or older. Do not include children 17 years or younger. Do not include anyone who is living somewhere else for more than 2 months, such as a college student living away or someone in the Armed Forces on deployment. How many people are currently living or staying at this address?

Number of people

[IF EXPERIMENT A = 6, ASK 1F, AS SPECIFIED]

1F1. The first question is about people at this address.

How many people are currently living or staying at this address?

Number of people

[If 1F1=1 then go to Q2. If 1F1>1 then ask 1F2. Keep 1F1 on the screen when ask 1F2.]
1F2. When you reported the number of people living or staying at this address, counting yourself, how many of them were 18 years or older?

Number of people

[IF 1F1=1F2 THEN SKIP TO 1F4, OTHERWISE GO TO 1F3. Keep 1F1 and 1F2 on the screen when ask 1F3]

1F3. When you reported the number of people living or staying at this address, how many of them were children 17 years or younger?

Number of people

[Keep 1F1, 1F2, and 1F3 on the screen when ask 1F4]

1F4. When you reported the number of people living or staying at this address, how many of them, if any, are currently living someplace else for more than two months, like a college student or someone in the Armed Forces on deployment?

Number of people

[Experiment 1]

Q2 through Q7

[Experiment A: Questions with Qualifications]

PROG NOTE: the respondent should receive items 8, 9, and 10 in the same condition as they received item 1.]

[IF EXPERIMENT A = 1 ASK 8A, 9A, and 10A]

8A. The next question is about your footwear.

How many pairs of shoes do you own?

Number of pairs of shoes

[Do not allow respondents to return to 8A]

9A. When you reported the pairs of shoes that you own, how many were:

Number of pairs of shoes

- 9A1. boots?.....
- 9A2. sneakers and/or athletic shoes?.....
- 9A3. bedroom slippers?.....
- 9A4. sandals?.....
- 9A5. other casual shoes?.....
- 9A6. dress shoes?.....

10A. The next question is about your coats and jackets.

How many coats and jackets do you own?

Number of coats and jackets

[IF EXPERIMENT A = 2 ASK 8B, 9B and 10B]

8B. The next question is about your footwear.

How many pairs of shoes do you own? For the purposes of this question, do not include boots, sneakers, athletic shoes, or bedroom slippers. Include sandals, other casual shoes, and dress shoes. If you do not own a pair of shoes (as we have defined them), enter "0."

Number of pairs of shoes

[Do not allow respondents to return to 8B]

9B. When you reported the pairs of shoes that you own, how many were:

Number of pairs of shoes

- 9B1. boots?.....
- 9B2. sneakers and/or athletic shoes?.....
- 9B3. bedroom slippers?.....
- 9B4. sandals?.....
- 9B5. other casual shoes?.....
- 9B6. dress shoes?.....

10B. The next question is about your coats and jackets.

How many coats and jackets do you own? For the purposes of this question, exclude coats and jackets for indoor use, such as sports coats, tailored jackets, blazers, or suits. Include

outerwear that is made from down or synthetic sources, such as polyester or acrylic. Do not include outerwear that is made from natural sources, such as fur, leather, wool, cotton or canvas. If you do not own a coat or jacket (as we have defined them), enter "0."

Number of coats and jackets

[IF EXPERIMENT A = 3 ASK 8C, 9C, and 10C]

8C. The next question is about your footwear.

For the purposes of this question, do not include boots, sneakers, athletic shoes, or bedroom slippers. Include sandals, other casual shoes, and dress shoes. If you do not own a pair of shoes (as we have defined them), enter "0." How many pairs of shoes do you own?

Number of pairs of shoes

[Do not allow respondents to return to 8C]

9C. When you reported the pairs of shoes that you own, how many were:

Number of pairs of shoes

- 9C1. boots?.....
- 9C2. sneakers and/or athletic shoes?.....
- 9C3. bedroom slippers?.....
- 9C4. sandals?.....
- 9C5. other casual shoes?.....
- 9C6. dress shoes?.....

10C. The next question is about your coats and jackets.

For the purposes of this question, exclude coats and jackets for indoor use, such as sports coats, tailored jackets, blazers, or suits. Include outerwear that is made from down or synthetic sources, such as polyester or acrylic. Do not include outerwear that is made from natural sources, such as fur, leather, wool, cotton or canvas. If you do not own a coat or jacket (as we have defined them), enter "0." How many coats and jackets do you own?

Number of coats and jackets

[IF EXPERIMENT A = 4 ASK 8D, 9D, and 10D]

8D. The next question is about your footwear.

How many pairs of shoes do you own? For the purposes of this question, do not include boots, sneakers, athletic shoes, or bedroom slippers. Include sandals, other casual shoes, and dress shoes. If you do not own a pair of shoes (as we have defined them), enter "0."

Number of pairs of shoes

[Do not allow respondents to return to 8D]

9D. When you reported the pairs of shoes that you own, how many were:

Number of pairs of shoes

- 9D1. boots?.....
- 9D2. sneakers and/or athletic shoes?.....
- 9D3. bedroom slippers?.....
- 9D4. sandals?.....
- 9D5. other casual shoes?.....
- 9D6. dress shoes?.....

10D. The next question is about your coats and jackets.

How many coats and jackets do you own? *For the purposes of this question, exclude coats and jackets for indoor use, such as sports coats, tailored jackets, blazers, or suits. Include outerwear that is made from down or synthetic sources, such as polyester or acrylic. Do not include outerwear that is made from natural sources, such as fur, leather, wool, cotton or canvas. If you do not own a coat or jacket (as we have defined them), enter "0."*

Number of coats and jackets

[IF EXPERIMENT A = 5 ASK 8E, 9E, and 10E]

8E. The next question is about your footwear.

For the purposes of this question, do not include boots, sneakers, athletic shoes, or bedroom slippers. Include sandals, other casual shoes, and dress shoes. If you do not own a pair of shoes (as we have defined them), enter "0." How many pairs of shoes do you own?

Number of pairs of shoes

[Do not allow respondents to return to 8E]

9E. When you reported the pairs of shoes that you own, how many were:

Number of pairs of shoes

- 9E1. boots?.....
- 9E2. sneakers and/or athletic shoes?.....
- 9E3. bedroom slippers?.....
- 9E4. sandals?.....
- 9E5. other casual shoes?.....
- 9E6. dress shoes?.....

10E. The next question is about your coats and jackets.

For the purposes of this question, exclude coats and jackets for indoor use, such as sports coats, tailored jackets, blazers, or suits. Include outerwear that is made from down or synthetic sources, such as polyester or acrylic. Do not include outerwear that is made from natural sources, such as fur, leather, wool, cotton or canvas. If you do not own a coat or jacket (as we have defined them), enter "0." How many coats and jackets do you own?

Number of coats and jackets

[IF EXPERIMENT A = 6 ASK 8F, 9F, AND 10F, AS SPECIFIED)

8F1. The next question is about your footwear.

How many pairs of shoes do you own?

Number of pairs of shoes

[If 8F1>0 ask 9F1. Keep 8F1 on screen when ask 9F1.]

9F1. When you reported the pairs of shoes that you own, how many pairs of boots, sneakers, athletic shoes, or bedroom slippers, if any, were included?

Number of pairs of shoes

[If 8F1>0 ask 9F2. Keep 8F1 and 9F1 on screen when ask 9F2. If 8F1 = 9F1 do not ask 9F2]

9F2. When you reported the pairs of shoes that you own, how many pairs of sandals, other casual shoes, or dress shoes, if any, were included?

Number of pairs of shoes

10F1. The next question is about your coats and jackets.

How many coats and jackets do you own?

Number of coats and jackets

[Keep 10F1 on the screen when ask 10F2]

10F2. When you reported the number of coats and jackets that you own, how many coats and jackets for indoor use, such as sports coats, tailored jackets, blazers, or suits, if any, were included?

Number of coats and jackets

[If 10F1=10F2 then do not ask 10F3. Keep 10F1 and 10F2 on the screen when ask 10F3]

10F3. When you reported the number of coats and jackets that you own, how many of them, if any, were outerwear made from down or synthetic sources, such as polyester or acrylic?

Number of coats and jackets

[Keep 10F1, 10F2, and 10F3 on the screen when ask 10F4]

10F4. When you reported the number of coats and jackets that you own, how many of them, if any, were outerwear made from natural sources, such as fur, leather, wool, cotton or canvas?

Number of coats and jackets

[Experiment 2]

Q11

[IF EXPERIMENT A = 1 ASK 12A AND 13A]

12A. Here are a few questions about work you may have performed.

Last week, how many hours, if any, did you work for either pay or profit?

Number of hours

[Do not allow respondents to return to 12A]

13A. When you reported the hours you worked for either pay or profit last week, how many were spent:

Number of hours

- 13A1. doing tasks required of your job or jobs?....._____
- 13A2. telecommuting or working from home?....._____
- 13A3. surfing the Web, working on personal matters, or socializing?....._____

[IF EXPERIMENT A = 2 ASK 12B AND 13B]

12B. Here are a few questions about work you may have performed.

Last week, how many hours, if any, did you work for either pay or profit? For the purposes of this question, do not count as work telecommuting or working from home. In addition, do not count time spent surfing the Web, working on personal matters or socializing as work. If you did not work for pay or profit last week (as we have defined it), enter "0."

Number of hours

[Do not allow respondents to return to 12B]

13B. When you reported the hours you worked for either pay or profit last week, how many were spent:

Number of hours

- 13B1. doing tasks required of your job or jobs?....._____
- 13B2. telecommuting or working from home?....._____
- 13B3. surfing the Web, working on personal matters, or socializing?....._____

[IF EXPERIMENT A = 3 ASK 12C AND 13C]

12C. Next are a few questions about work you may have performed.

For the purposes of this question, do not count as work telecommuting or working from home. In addition, do not count time spent surfing the Web, working on personal matters, or socializing as work. If you did not work for pay or profit last week (as we have defined it), enter "0." Last week, how many hours, if any, did you work for either pay or profit?

Number of hours

[Do not allow respondents to return to 12 C]

13C. When you reported the hours you worked for either pay or profit last week, how many were spent:

Number of hours

13C1. doing tasks required of your job or jobs?... _____

13C2. telecommuting or working from home?..... _____

13C3. surfing the Web, working on personal matters, or socializing?..... _____

[IF EXPERIMENT A = 4 ASK 12D AND 13D]

12D. Here are a few questions about work you may have performed.

Last week, how many hours, if any, did you work for either pay or profit?
For the purposes of this question, do not count as work telecommuting or working from home. In addition, do not count time spent surfing the Web, working on personal matters, or socializing. If you did not work for pay or profit last week (as we have defined it), enter "0."

Number of hours

[Do not allow respondents to return to 12D]

13D. When you reported the hours you worked for either pay or profit last week, how many were spent:

Number of hours

13D1. doing tasks required of your job or jobs?... _____

13D2. telecommuting or working from home?..... _____

13D3. surfing the Web, working on personal matters, or socializing?..... _____

[IF EXPERIMENT A = 5 ASK 12E, 13E, and 14E]

12E. Here are a few questions about work you may have performed.

For the purposes of this question, do not count as work telecommuting or working from home. In addition, do not count time spent surfing the Web, working on personal matters, or socializing as work. If you did not work for pay or profit last week (as we have defined it), enter "0." Last week, how many hours, if any, did you work for either pay or profit?

Number of hours

[Do not allow respondents to return to 12E]

13E. When you reported the hours you worked for either pay or profit last week, how many were spent:

Number of hours

- 13E1. doing tasks required of your job or jobs?... _____
- 13E2. telecommuting or working from home?..... _____
- 13E3. surfing the Web, working on personal matters, or socializing?..... _____

[IF EXPERIMENT A = 6 ASK 12F and 13F]

12F1. Here are a few questions about work you may have performed.

Last week, how many hours, if any, did you work for either pay or profit?

Number of hours

[If 12F1>0 ask 13F1. Keep 12F1 on the screen when ask 13F1]

13F1 When you reported the hours you worked last week, how many hours, if any, were spent telecommuting or working from home?

Number of hours

[If 12F1>0 ask 13F2. Keep 12F1 and 13F1 on screen when ask 13F2]

13F2 When you reported the hours you worked last week, how many hours, if any, were spent surfing the Web, working on personal matters, or socializing?

Number of hours

[Experiment 3: Grouping of Response Options]

Q14 through Q16

[Experiment A: Questions with Qualifications]

PROG. NOTE: the respondent should receive items 17 and 18 in the same condition as they received item 1.]

[IF EXPERIMENT A = 1 ASK 17A and 18A]

17A. Now, a question about times when you may have been away from home for personal reasons.

In the past year, how many times, if any, were you away from home on a trip?

Number of trips

18A. The next question is about furniture purchases you may have made.

In the past year, how many furniture purchases, if any, did you make?

Number of purchases

[IF EXPERIMENT A = 2 ASK 17B and 18B]

17B. Now, a question about times when you may have been away from home for personal reasons.

In the past year, how many times, if any, were you away from home on a trip? For the purposes of this question, only report times when you were away from home for 3 or more nights in a row. Exclude all business trips. If you were not away from home (as we have defined it) in the past year, enter "0."

Number of trips

18B. The next question is about furniture purchases you may have made.

In the past year, how many furniture purchases, if any, did you make? For the purposes of this question, do not include items such as TVs, radios, lamps and lighting fixtures, outdoor furniture, infant's furniture, or appliances as furniture. Include items such as tables, chairs, footstools, or sofas as furniture. If you did not purchase any furniture (as we have defined it) in the past year, enter "0".

Number of purchases

[IF EXPERIMENT A = 3 ASK 17C and 18C]

17C. Now, a question about times when you may have been away from home for personal reasons.

For the purposes of this question, only report times when you were away from home for 3 or more nights in a row. Exclude all business trips. If you were not away from home (as we have defined it) in the past year, enter "0." In the past year, how many times, if any, were you away from home on a trip?

Number of trips

18C. The next question is about furniture purchases you may have made.

For the purposes of this question, do not include items such as TVs, radios, lamps and lighting fixtures, outdoor furniture, infant's furniture, or appliances. Include items such as tables, chairs, footstools, or sofas as furniture. If you did not purchase any furniture (as we have defined it) in the past year, enter "0". In the past year, how many furniture purchases, if any, did you make?

Number of purchases

[IF EXPERIMENT A = 4 ASK 17D and 18D]

17D. Now, a question about times when you may have been away from home for personal reasons.

In the past year, how many times, if any, were you away from home on a trip? *For the purposes of this question, only report times when you were away from home for 3 or more nights in a row. Exclude all business trips. If you were not away from home (as we have defined it) in the past year, enter "0."*

Number of trips

18D. The next question is about furniture purchases you may have made.

In the past year, how many furniture purchases, if any, did you make? ? *For the purposes of this question, do not include items such as TVs, radios, lamps and lighting fixtures, outdoor furniture, infant's furniture, or appliances. Include items such as tables, chairs, footstools, or sofas as furniture. If you did not purchase any furniture (as we have defined it) in the past year, enter "0".*

Number of purchases

[IF EXPERIMENT A = 5 ASK 17E AND 18E]

17E. Now, a question about times when you may have been away from home for personal reasons.

For the purposes of this question, only report times when you were away from home for 3 or more nights in a row. Exclude all business trips. If you were not away from home (as we have defined it) in the past year, enter "0." In the past year, how many times, if any, were you away from home on a trip?

Number of trips

18E. The next question is about furniture purchases you may have made.

For the purposes of this question, do not include items such as TVs, radios, lamps and lighting fixtures, outdoor furniture, infant's furniture, or appliances. Include items such as tables, chairs, footstools, or sofas as furniture. If you did not purchase any furniture (as we have defined it) in the past year, enter "0". In the past year, how many furniture purchases, if any, did you make?

Number of purchases

[IF EXPERIMENT A = 6 ASK 17F, AS SPECIFIED]

17F1. Now, a question about times when you may have been away from home for personal reasons.

In the past year, how many times, if any, were you away from home on a trip?

Number of trips

[If 17F1>0 ask 17F2. Keep 17F1 on screen when ask 17F2. If 17F1=0, go to 18]

17F2 When you reported the number of times you were away from home, how many of them, if any, were for 3 or more nights?

Number of trips

[Keep 17F2 on screen when ask 17F3]

17F3 When you reported the number of times you were away from home for 3 or more nights, how many of them, if any, were for business?

Number of trips

18F1. The next question is about furniture purchases you may have made.

In the past year, how many furniture purchases, if any, did you make?

Number of purchases

[If 18F1>0 then ask 18F2; Keep 18F1 on screen when ask 18F1]

18F2. When you reported purchasing furniture in the past year, how many of these purchases, if any, included TVs, radios, lamps and lighting fixtures, outdoor furniture, and infants' furniture or appliances?

Number of purchases

[If 18F1>0 then ask 18F3, unless 18F1=18F2 then do not ask 18F3; Keep 18F1 and 18F2 on screen when ask 18F3]

18F3. When you reported purchasing furniture in the past year, how many of these purchases, if any, included tables, chairs, footstools, sofas, and so on?

Number of purchases or expenses

[Experiment 4]

Q19 through Q25

[Experiment A: Questions with Qualifications]

PROG. NOTE: the respondent should receive items 26 and 27 in the same condition as they received item 1.

[IF EXPERIMENT A = 1 ASK 26A and 27A]

26A. The next question is about the bedrooms in this house, apartment, or mobile home.

How many bedrooms are in this house, apartment, or mobile home?

Number of bedrooms

27A. The next question is about the other rooms in this house, apartment, or mobile home.

How many other separate rooms are in this house, apartment, or mobile home?

Number of rooms

[IF EXPERIMENT A = 2 ASK 26B and 27B]

26B. The next question is about the bedrooms in this house, apartment, or mobile home.

How many bedrooms are in this house, apartment, or mobile home? For the purposes of this question, only include those bedrooms located on the main floor of this house, apartment, or mobile home. If you do not have any bedrooms (as we have defined them), enter '0.' If this is an efficiency/studio apartment, enter "0".

Number of bedrooms

27B. The next question is about the other rooms in this house, apartment, or mobile home.

How many other separate rooms are in this house, apartment, or mobile home? For the purposes of this question, rooms must be separated by built-in archways or walls that extend out at least 6 inches and go from floor to ceiling. Exclude bedrooms, bathrooms, laundry rooms, porches, balconies, foyers, halls, or unfinished basements.

Number of rooms

[IF EXPERIMENT A = 3 ASK 26C and 27C]

26C. The next question is about the bedrooms in this house, apartment, or mobile home.

For the purposes of this question, only include those bedrooms located on the main floor of this house, apartment, or mobile home. If you do not have any bedrooms (as we have defined them), enter '0.' If this is an efficiency/studio apartment, enter "0." How many bedrooms are in this house, apartment, or mobile home?

Number of bedrooms

27C. The next question is about the other rooms in this house, apartment, or mobile home.

For the purposes of this question, rooms must be separated by built-in archways or walls that extend out at least 6 inches and go from floor to ceiling. Exclude bedrooms, bathrooms, laundry rooms, porches, balconies, foyers, halls, or unfinished basements. How many other separate rooms are in this house, apartment, or mobile home?

Number of rooms

[IF EXPERIMENT A = 4 ASK 26D and 27D]

26D. The next question is about the bedrooms in this house, apartment, or mobile home.

How many bedrooms are in this house, apartment, or mobile home? *For the purposes of this question, only include those bedrooms located on the main floor of this house, apartment, or mobile home. If you do not have any bedrooms (as we have defined them), enter '0.' If this is an efficiency/studio apartment, enter "0".*

Number of bedrooms

27D. The next question is about the other rooms in this house, apartment, or mobile home.

How many other separate rooms are in this house, apartment, or mobile home? *For the purposes of this question, rooms must be separated by built-in archways or walls that extend out at least 6 inches and go from floor to ceiling. Exclude bedrooms, bathrooms, laundry rooms, porches, balconies, foyers, halls, or unfinished basements.*

Number of rooms

[IF EXPERIMENT A = 5 ASK 26E and 27E]

26E. The next question is about the bedrooms in this house, apartment, or mobile home.

For the purposes of this question, only include those bedrooms located on the main floor of this house, apartment, or mobile home. If you do not have any bedrooms (as we have defined them), enter '0.' If this is an efficiency/studio apartment, enter "0". How many bedrooms are in this house, apartment or mobile home?

Number of bedrooms

27E. The next question is about the other rooms in this house, apartment, or mobile home.

For the purposes of this question, rooms must be separated by built-in archways or walls that extend out at least 6 inches and go from floor to ceiling. Exclude bedrooms, bathrooms, laundry rooms, porches, balconies, foyers, halls, or unfinished basements. How many other separate rooms are in this house, apartment, or mobile home?

Number of rooms

[IF EXPERIMENT A = 6 ASK 26F and 27F, AS SPECIFIED]

26F1. The next question is about bedrooms in this house, apartment, or mobile home.

How many bedrooms are in this house, apartment, or mobile home?

Number of bedrooms

[Keep 27F2 on the screen when ask 27F3]

26F2 When you reported the number of bedrooms in this house, apartment or mobile home, how many bedrooms located on the main floor, if any, were included?

Number of bedrooms

[If 26F1=0 or 26F1=1, then ask 23F3]

26F3. Is this is an efficiency or studio apartment?

- 1 Yes
- 2 No

27F1 The next question is about the other rooms in this house, apartment, or mobile home.

How many other rooms are in this house, apartment, or mobile home that are separated by built-in archways or have walls that extend out at least 6 inches and go from floor to ceiling?

Number of rooms

[If 27F1>0 ask 27F2. Keep 27F1 on the screen when ask 27F2]

27F2 When you reported the number of other rooms in this house, apartment, or mobile home, how many bedrooms, bathrooms, laundry rooms, porches, balconies, foyers, halls, or unfinished basements, if any, were included?

Number of rooms

A.2 Experiment 2

2010 Channel Experiment Questionnaire– 2/22//2010

PROGRAMMING NOTES:

1. Two Questionnaire Versions: IVR and Web
 2. Capture timestamp at every question of main survey (1 through 25). They need to be programmed for both versions.
 3. For IVR Survey: Skip the question if respondent enters an invalid key 3 times.
 4. For IVR Survey: If 10 second period elapses without answering prompt with “We would be very grateful if you would be willing to provide your best answer, even if you are not completely sure. Please enter your response followed by the # key. If you would like me to repeat the question, press the star (*) key”. If additional 10 second period elapses without answering the prompt “Please enter your response followed by the # key. If you would like me to repeat the question, press the star (*) key. But if you prefer to skip this question, press the pound (#) key.”
 5. Do not display “Don’t know” and “Refused” options for IVR and Web surveys. Respondents for IVR and Web surveys can skip to next question without answering a question as follows:
 - a. For WEB survey: By pressing “next” button;
 - b. For IVR survey: By pressing “#” key “skip”
 6. For IVR Survey respondent can have question repeated by pressing the “*” key
 7. If respondent skips a question without answering, use the following soft prompts:
 - a. WEB: We noticed that you did not answer the question below. We would be grateful if you would provide your best answer, even if you’re not completely sure. But if you would prefer to skip this question, you can click “Next.”
 - b. IVR: We noticed that you did not answer the previous question. We would be grateful if you would provide your best answer, even if you’re not completely sure. Please enter your response followed by the # key. If you would like me to repeat the question, press the star (*) key. But if you would prefer to skip this question, press the pound (#) key.”
 8. If response is out of range, use the following soft prompt:
 - a. WEB: This response seems high. If this is a mistake, please correct. But if this number is correct, click “Next” to continue.
 - b. IVR: This response seems high. If this is a mistake, press the star (*) key to correct. But if this number is correct, press the pound (#) key to continue.
-

Redline_Channel Experiment_Questionnaire

[TIME STAMP]

Hello, my name is [INTERVIEWER'S NAME] and I'm calling on behalf of the University of Maryland. You have been randomly selected to participate in a brief study about health practices and lifestyles. In order to determine the person I should speak to, may I please speak to the adult, age 18 or older, who most recently celebrated a birthday?

[ONCE RESP IS ON THE PHONE:]

Hello, my name is [INTERVIEWER'S NAME] and I'm calling on behalf of the University of Maryland. You have been randomly selected to participate in a brief study about health practices and lifestyles. Your participation is voluntary, but critical for the success of the study. All of your responses will be kept confidential. You may skip any questions you don't want to answer. The survey will take about 5 minutes. May we begin?

SCREENER [CATI]

First, I'm going to ask you a few background questions.

1. What is your gender? [NOTE: ASK ONLY IF NECESSARY]

- 1 Male
- 2 Female

2. In what year were you born?

[RECORD YEAR AS 4 DIGITS; 1900-1992; 9999=REFUSED]

3. Are you of Hispanic origin or descent, such as Mexican, Puerto Rican, Cuban, or other Spanish background?

- 1 Yes
- 2 No
- 8 DON'T KNOW
- 9 REFUSED

4. What is your primary race? Do you consider yourself to be...

- 1 Black or African-American
- 2 White
- 3 Asian or Pacific Islander
- 4 Native American, American Indian or Alaskan Native
- 5 Some other race

Redline_Channel Experiment_Questionnaire

- 8 DON'T KNOW
- 9 REFUSED

5. What is the highest degree or level of school you have completed?

- 1 Less than high school
- 2 High school graduate – high school diploma or the equivalent
- 3 Some college / Associate degree
- 4 Bachelor's degree
- 5 Master's degree
- 6 Doctorate / Professional degree
- 8 DON'T KNOW
- 9 REFUSED

6. Do you have access to the Internet for personal use?

- 1 Yes
- 2 No
- 8 DON'T KNOW
- 9 REFUSED

EXP1: MODE ASSIGNMENT MODULE:

[CATI: IF Q.6=2-9 THANK AND TERMINATE]

[CATI: RANDOMLY ASSIGN TO MODE]

MODE:

- 1 WEB SKIP TO W1
- 2 IVR SKIP TO I1

EXP1: SPLIT RESPONDENTS ACROSS MODES

II. MODE=2: IVR

Thank you for answering those background questions. Now I'd like to switch you over to an automated response system for some questions about your health practices and lifestyle. We have designed the system to be easy to use and hope it will speed your response. The rest of the interview will take about 10 minutes. If you complete the survey, you will receive \$5.00 in the mail.

- 1. Agreed to participate in automated interview
- 2. Refusal to participate → **SKIP TO Q.XX = 10**
- 3. Schedule callback

Redline_Channel Experiment_Questionnaire

I1a. In order to send you \$5.00, I need to ask your name and mailing address.

1 Continue

2. Do not want to receive \$5

ENTER ADDRESS INFORMATION BELOW:

Record First name: _____

Record Last name: _____

Record Street Address: _____

Record City: _____

Record State: _____

Record Zip Code: _____

I1b. What phone number did I reach you at?

1. Gave response (Record 10 DIGIT # _____)
2. DON'T KNOW/REFUSED

I2. I'm going to switch you over to the automated system now. You can use the keypad of your phone to enter your responses followed by the # (pound) key. Press the * (star) key if you would like to repeat a question and the # (pound) key to skip a question. As mentioned before, this will take about 10 minutes. I appreciate your participation in this important research. Please hold while I transfer you.

[INTERVIEWER: ENTER "1" BELOW TO TRANSFER THE CALL TO IVR SYSTEM]

- 1> Transferred call to IVR system → **SKIP TO Q.XX (= 9)**
- 2> Refusal (THANK & TERMINATE → **SKIP TO Q.XX (= 10)**)
- 3> Schedule callback

SKIP TO INTERVIEW CLASSIFICATION QUESTION XX COUNT IT AS COMPLETED INTERVIEW

IF MODE=2: WEB

W1. Thank you for answering those background questions. Based on your responses, you have qualified to participate in our Web Survey. We would like to send you a link to a short Web questionnaire about your health practices and lifestyle that you can complete in your own time. If you complete the survey online, you will receive a \$10 incentive at the end of the survey.

- 1> Yes, agreed to participate
- 2> Refused (THANK & TERMINATE) → **SKIP TO Q.XX (= 4)**
- 3> Schedule callback

Redline_Channel Experiment_Questionnaire

W2. In order for us to send you information about how to access the survey, we would like your email address. Your email address will not be sold or shared with anyone else. What is the best e-mail address to reach you?

[READ IF NECESSARY:] Your email address will be kept completely confidential. It will not be sold, given, or shared with anyone else. Your email address will only be used for the purpose of this research study.

[INTERVIEWER NOTE: Example of Email address: jsmith@aol.com]

- 1> PROVIDED EMAIL ADDRESS (RECORD ADDRESS: _____)
- 2> DOES NOT HAVE EMAIL → **SKIP TO W5**
- 3> REFUSED EMAIL ADDRESS → **SKIP TO W5**

W3 Let me see if I got that right. You told me that your email address is [EMAIL ADDRESS FROM W2]. Is that right?

[INTERVIEWER: EXTREMELY IMPORTANT TO ENTER THE EMAIL ADDRESS ACCURATELY, Example of Email address: jsmith@aol.com]

- 1> Correct
- 2> Enter Correction

W4. You will receive an e-mail from us shortly with the link to the questionnaire. If you would like I can also read the web address for the questionnaire to you and provide you with your unique login ID now (IF RESPONDENT WOULD LIKE IT NOW: The website address is www.opinionport.com/mdlifestyles and your login ID is your ten-digit phone number, that is xxxxxxxxxx.). The sender of the email will be mdlifestyles@srbj.com and the subject line will read "University of Maryland Health and Lifestyles Survey. Thank you in advance for filling it out; your response is very important to our study. → **SKIP TO Q.XX (= 3)**

W5 [IF W2=3 I understand your hesitation to provide your email address.']. If you would like I can read the web address for the questionnaire to you instead, or I can fax it to you. The website address is www.opinionport.com/mdlifestyles and your unique login ID password is your ten-digit phone number, that is xxxxxxxxxx. As mentioned before, if you complete the survey online, you will receive \$10 in the mail.]

- 1> Requested fax (Record Fax# _____) → **ASK W6 THEN SKIP TO Q.XX(=7)**
- 2> Read link and login password → **ASK W6 THEN SKIP TO Q.XX (= 8)**
- 3> Refused to participate → **SKIP TO Q.XX (= 5 OR 6)**

Redline_Channel Experiment_Questionnaire

W6. Great, I need to verify your name and mailing address in order to send you \$10.00

1. Continue
2. Do not want to receive \$10 (SKIP TO QXX, CODE QXX BASED ON W5)

ENTER ADDRESS INFORMATION BELOW:

Record First name: _____
Record Last name: _____
Record Street Address: _____
Record City: _____
Record State: _____
Record Zip Code: _____

EVERYBODY ANSWERING W6 SKIPS TO Q.XX

MAIN SURVEY

**[IVR AND WEB QUESTIONNAIRES START FROM HERE]
[BEGINNING MAIN SURVEY TIME STAMP]**

[WEB INTRODUCTION]

Welcome to the University of Maryland Health and Lifestyles Survey. Thank you participating in this study. The survey will take about 10 minutes. As a token of our appreciation, you will receive \$10 upon completion of this survey.

Please enter your 10-digit phone number to start the survey. If you have already started the survey, you will be brought back to where you left off. If at any time you would like to skip a question, please press the “Next” button.

As you move through the survey, please use only the Next button at the bottom of the screen. Please do not use the Back or Forward buttons on your browser.

_____ Enter your 10-digit phone number with no dashes, parentheses or spaces to start the survey.

SKIP TO Q.1A.

[IVR INTRODUCTION]

1. Thank you for participating in the University of Maryland's Health and Lifestyles Survey. You can enter your responses by using the keypad of your phone followed by pressing the # (pound) key. If at any time you would like to skip a question, press the # (pound) key. If you would like to have a question repeated, press the * (star) key.

IVR VERSION ONLY

M1 Before we begin, may I please have the ten digit phone number including area code that we reached you at? Again, please use the keypad of your phone to enter your answer followed by pressing the pound key

_____ **(ACCEPT 10 DIGITS)**

SKIP TO Q.1A.

[Experiment A: Questions with Lengthy Clarifying Information

This is a one-factor experiment. Present one item per screen with labeled numeric entry box. There are 12 items in this experiment, blocks of which are interspersed with buffer questions. The twelve items are: 1, 6,7,8,12,13,14,18,19,20,24, and 25. Once a respondent is randomly assigned to one of the six conditions below, the respondent should receive all ten items in that one condition.

Experiment A

1. **One Question, Qualifications After, IVR**
2. **One Question, Qualifications After, Web**
3. **One Question, Qualifications Before, IVR**
4. **One Question, Qualifications Before, Web**
5. **Multiple Questions with Qualifications Incorporated, IVR**
6. **Multiple Questions with Qualifications Incorporated, Web]**

[IF EXPERIMENT A = 1 ASK 1A in IVR]
[IF EXPERIMENT A = 2 ASK 1A in Web]

1A. The first question is about the people at your home address.

How many people are currently living or staying at your home address? Do not forget to count yourself. For the purposes of this question, a person is defined as someone 18 years or older. Do not include children 17 years or younger. Do not include anyone who is living somewhere else for more than 2 months, such as a college student living away or someone in the Armed Forces on deployment.

Redline_Channel Experiment_Questionnaire

Number of people

[IF EXPERIMENT A = 3 ASK 1B in IVR]

[IF EXPERIMENT A = 4 ASK 1B in Web]

1B. The first question is about the people at your home address.

Do not forget to count yourself. For the purposes of this question, a person is defined as someone 18 years or older. Do not include children 17 years or younger. Do not include anyone who is living somewhere else for more than 2 months, such as a college student living away or someone in the Armed Forces on deployment. How many people are currently living or staying at your home address?

Number of people

[IF EXPERIMENT A = 5, ASK 1C in IVR]

[IF EXPERIMENT A = 6, ASK 1C in Web]

1C1. The first question is about people at your home address.

How many people are currently living or staying at your home address?

Number of people

[If 1C1=1 then go to Q2. If 1C1>1 then ask 1C2. Keep 1C1 on the screen when ask 1C2.]

1C2. When you reported the number of people living or staying at your home address, counting yourself, how many of them were 18 years or older?

Number of people

[IF 1C1=1C2 THEN SKIP TO 1C4, OTHERWISE GO TO 1C3. Keep 1C1 and 1C2 on the screen when ask 1C3]

1C3. When you reported the number of people living or staying at your home address, how many of them were children 17 years or younger?

Number of people

[Keep 1C1, 1C2, and 1C3 on the screen when ask 1C4]

1C4. When you reported the number of people living or staying at your home address, how many of them, if any, are currently living someplace else for more than two months, like a college student or someone in the Armed Forces on deployment?

Number of people

Experiment B: Examples

This is a one-factor experiment. Present one item per screen.

Experiment B

- 1 Qualifications After, IVR**
- 2 Qualifications After, Web**
- 3 Qualifications Before, IVR**
- 4 Qualifications Before, Web**

[IF EXPERIMENT B = 1 ASK 2A through 5A in IVR]

[IF EXPERIMENT B = 2 ASK 2A through 5A in Web]

[PROGRAM NOTE: RANDOMIZE QUESTION ORDER]

[PROG NOTE: the respondent should receive items 2 through 5 in the same condition as they received item 1, unless they had received conditions 5 or 6, in which case, respondents who had received condition 5 need to be randomly re-assigned condition 1 or 3. Those receiving condition 6, need to be randomly re-assigned to condition 2 or 4.]

Next are some questions about foods you may usually eat or drink.

2A. How many servings of meat do you typically eat each week? For the purposes of this question, meat includes beef, pork, poultry, and other meat.

[RECORD NUMBER 0-999] servings

3A. How many servings of dairy products do you typically eat each week? For the purposes of this question, dairy products include milk, cheese, yogurt, and other dairy products

[RECORD NUMBER 0-999] servings

4A. How many servings of grain products do you typically eat each week? For the purposes of this question, grain products include bread, pasta, rice, and other grain products.

[RECORD NUMBER 0-999] servings

Redline_Channel Experiment_Questionnaire

5A. How many servings of fruits and vegetables do you typically eat each week? For the purposes of this question, fruits and vegetables include apples, bananas, lettuce, and other fruits and vegetables.

[RECORD NUMBER 0-999] servings

[IF EXPERIMENT B = 3 ASK 2B through 5B in IVR]
[IF EXPERIMENT B = 4 ASK 2B through 5B in Web]
[PROGRAM NOTE: RANDOMIZE QUESTION ORDER]

Next are some questions about foods you may usually eat or drink.

2B. For the purposes of this question, meat includes beef, pork, poultry, and other meat. How many servings of meat do you typically eat each week?

[RECORD NUMBER 0-999] servings

3B. For the purposes of this question, dairy products include milk, cheese, yogurt, and other dairy products. How many servings of dairy products do you typically eat each week?

[RECORD NUMBER 0-999] servings

4B. For the purposes of this question, grain products include bread, pasta, rice, and other grain products. How many servings of grain products do you typically eat each week?

[RECORD NUMBER 0-999] servings

5B. For the purposes of this question, fruits and vegetables include apples, bananas, lettuce and other fruits and vegetables. How many servings of fruits and vegetables do you typically eat each week?

[RECORD NUMBER 0-999] servings

[Experiment A: Questions with Qualifications

PROG NOTE: the respondent should receive items 6, 7, and 8 in the same condition as they received item 1.]

[IF EXPERIMENT A = 1 ASK 6A, 7A and 8A in IVR]
[IF EXPERIMENT A = 2 ASK 6A, 7A and 8A in Web]

6A. The next question is about your footwear.

How many pairs of shoes do you own? For the purposes of this question, do not include boots, sneakers, athletic shoes, or bedroom slippers. Include sandals, other casual shoes, and dress shoes. If you do not own a pair of shoes (as we have defined them), enter "0."

Number of pairs of shoes

Redline_Channel Experiment_Questionnaire

[Do not allow respondents to return to 6A]

[ASK ONLY IF 6a>0]

7A. When you reported the pairs of shoes that you own, how many were

Number of pairs of shoes

- 7A1. boots?..... _____
- 7A2. sneakers and/or athletic shoes?..... _____
- 7A3. bedroom slippers?..... _____
- 7A4. sandals?..... _____
- 7A5. other casual shoes?..... _____
- 7A6. dress shoes?..... _____

8A. The next question is about your coats and jackets.

How many coats and jackets do you own? For the purposes of this question, exclude coats and jackets for indoor use, such as sports coats, tailored jackets, blazers, or suits. Include outerwear that is made from down or synthetic sources, such as polyester or acrylic. Do not include outerwear that is made from natural sources, such as fur, leather, wool, cotton or canvas. If you do not own a coat or jacket (as we have defined them), enter "0."

Number of coats and jackets

[IF EXPERIMENT A = 3 ASK 6B, 7B, and 8B]

[IF EXPERIMENT A = 4 ASK 6B, 7B, and 8B]

6B. The next question is about your footwear.

For the purposes of this question, do not include boots, sneakers, athletic shoes, or bedroom slippers. Include sandals, other casual shoes, and dress shoes. If you do not own a pair of shoes (as we have defined them), enter "0." How many pairs of shoes do you own?

Number of pairs of shoes

[Do not allow respondents to return to 6B]

[ASK ONLY IF 6b>0]

7B. When you reported the pairs of shoes that you own, how many were:

Redline_Channel Experiment_Questionnaire

Number of pairs of shoes

- 7B1. boots?.....
- 7B2. sneakers and/or athletic shoes?.....
- 7B3. bedroom slippers?.....
- 7B4. sandals?.....
- 7B5. other casual shoes?.....
- 7B6. dress shoes?.....

8B. The next question is about your coats and jackets.

For the purposes of this question, exclude coats and jackets for indoor use, such as sports coats, tailored jackets, blazers, or suits. Include outerwear that is made from down or synthetic sources, such as polyester or acrylic. Do not include outerwear that is made from natural sources, such as fur, leather, wool, cotton or canvas. If you do not own a coat or jacket (as we have defined them), enter "0." How many coats and jackets do you own?

Number of coats and jackets

[IF EXPERIMENT A = 5 ASK 6C, 7C, AND 8C in IVR]
[IF EXPERIMENT A = 6 ASK 6C, 7C, AND 8C in Web]

6C1. The next question is about your footwear.

How many pairs of shoes do you own?

Number of pairs of shoes

[If 6C1>0 ask 7C1. Keep 6C1 on screen when ask 7C1.]

7C1. When you reported the pairs of shoes that you own, how many pairs of boots, sneakers, athletic shoes, or bedroom slippers, if any, were included?

Number of pairs of shoes

[If 6C1>0 ask 7C2. Keep 6C1 and 7C1 on screen when ask 7C2. If 6C1 = 7C1 do not ask 7C2]

7C2. When you reported the pairs of shoes that you own, how many pairs of sandals, other casual shoes, or dress shoes, if any, were included?

Number of pairs of shoes

8C1. The next question is about your coats and jackets.

Redline_Channel Experiment_Questionnaire

How many coats and jackets do you own?

Number of coats and jackets

If 8C1=0, SKIP TO Q9

[Keep 8C1 on the screen when ask 8C2]

8C2. When you reported the number of coats and jackets that you own, how many coats and jackets for indoor use, such as sports coats, tailored jackets, blazers, or suits, if any, were included?

Number of coats and jackets

[If 8C1=8C2 then do not ask 8C3 or 8C4. Keep 8C1 and 8C2 on the screen when ask 8C3]

8C3. When you reported the number of coats and jackets that you own, how many of them, if any, were outerwear made from down or synthetic sources, such as polyester or acrylic?

Number of coats and jackets

[Keep 8C1, 8C2, and 8C3 on the screen when ask 8C4]

8C4 When you reported the number of coats and jackets that you own, how many of them, if any, were outerwear made from natural sources, such as fur, leather, wool, cotton or canvas?

Number of coats and jackets

[Buffer Question: 9, 10, 11]

Now for some questions about your lifestyle.

9. Many people say they have less time these days to do volunteer work. What about you, were you able to devote any time to volunteer work in the last 12 months?

- 1 Yes
- 2 No

10. If you were selected to serve on a jury, would you be happy to do it or would you rather not serve?

- 1 Yes, would serve
- 2 No, would rather not serve

Redline_Channel Experiment_Questionnaire

11. Many people are finding it difficult to make contributions to church or charity as much as they used to. How about you? Were you able to contribute any money to church or charity in the last 12 months?

- 1 Yes
- 2 No

[Experiment A: Questions with Qualifications

PROG NOTE: the respondent should receive items 12, 13, and 14 in the same condition as they received item 1.]

[IF EXPERIMENT A = 1 ASK 12A, 13A and 14A in IVR]
[IF EXPERIMENT A = 2 ASK 12A, 13A and 14A in Web]

12A. Now for a question about communications you may have had .

In the past week, how many telephone calls did you make or receive? For the purposes of this question, include telephone calls you made or received at home. Do not include telephone calls you made or received away from home, for example, while working, commuting, or shopping. If you did not make or receive any telephone calls (as we have defined them) in the past week, enter "0."

Number of calls

13A. Here are a few questions about work you may have performed.

Last week, how many hours, if any, did you work for either pay or profit? For the purposes of this question, do not count as work telecommuting or working from home. In addition, do not count time spent surfing the Web, working on personal matters or socializing at work. If you did not work for pay or profit last week (as we have defined it), enter "0."

Number of hours

[Do not allow respondents to return to 13A]
[ASK ONLY IF 13A>0]

14A. When you reported the hours you worked for either pay or profit last week, how many were spent:

Redline_Channel Experiment_Questionnaire

Number of hours (RANGE 0 – 70)

- 14A1. doing tasks required of your job or jobs?... _____
- 14A2. telecommuting or working from home?..... _____
- 14A3. surfing the Web, working on personal matters, or socializing at work?..... _____

[IF EXPERIMENT A = 3 ASK 12B, 13B, and 14B]
[IF EXPERIMENT A = 4 ASK 12B, 13B, and 14B]

12B. Now for a question about communications you may have had..

For the purposes of this question, include telephone calls you made or received at home. Do not include telephone calls you made or received away from home, for example, while working, commuting, or shopping. If you did not make or receive any telephone calls (as we have defined them) in the past week, enter "0." In the past week, how many telephone calls did you make or receive?

Number of calls

13B. Next are a few questions about work you may have performed.

For the purposes of this question, do not count as work telecommuting or working from home. In addition, do not count time spent surfing the Web, working on personal matters, or socializing at work. If you did not work for pay or profit last week (as we have defined it), enter "0." Last week, how many hours, if any, did you work for either pay or profit?

Number of hours

[Do not allow respondents to return to 13B]
[ASK ONLY IF 13B>0]

14B. When you reported the hours you worked for either pay or profit last week, how many were spent:

Number of hours

- 14B1. doing tasks required of your job or jobs?... _____
- 14B2. telecommuting or working from home?..... _____
- 14B3. surfing the Web, working on personal matters, or socializing at work?..... _____

Redline_Channel Experiment_Questionnaire

[IF EXPERIMENT A = 5 ASK 12C, 13C, and 14C]

[IF EXPERIMENT A = 6 ASK 12C, 13C, and 14C]

12C1. Now for a question about communications you may have had..

In the past week, how many telephone calls did you make or receive?

Number of calls

[If 12C1>0 then ask 12C2. Keep 12C1 on screen when ask 12C2.]

12C2. When you reported the number telephone calls you made or received last week, how many of them , if any, included telephone calls you made or received at home?.

Number of ca.lls

[If 12C1>0 then ask 12C3, unless 12C1=12C2. Keep 12C1 and 12C2 on screen when ask 12C3.]

12C3. When you reported the number telephone calls you made or received last week, how many of them , if any, included telephone calls you made or received away from home, for example, while working, commuting, or shopping?

Number of ca.lls

13C1. Here are a few questions about work you may have performed.

Last week, how many hours, if any, did you work for either pay or profit?

Number of hours

[If 13C1>0 ask 14C1. Keep 13C1 on the screen when ask 14C1]

14C1 When you reported the hours you worked last week, how many hours, if any, were spent telecommuting or working from home?

Number of hours

Redline_Channel Experiment_Questionnaire

[If 13C1>0 ask 14C2. Keep 13C1 and 14C1 on screen when ask 14C2]

14C2 When you reported the hours you worked last week, how many hours, if any, were spent surfing the Web, working on personal matters, or socializing at work?

Number of hours (RANGE 0 – 70)

[Buffer Questions: 15, 16, AND 17]

15 FOR IVR:

Please describe your current EMPLOYMENT STATUS. Are you [INSERT ITEM]?

- a Employed for wages
- b Self-employed

- 1. YES
- 2. NO

[IF Q15a = 1 or Q15b = 1 then ASK 16, OTHERWISE SKIP TO Q17]

15 FOR WEB

Please describe your current EMPLOYMENT STATUS. Are you:

	Yes	No
a. Employed for wages		
b. Self-employed		

[IF Q15a = yes or Q15b = yes then ASK 16, OTHERWISE SKIP TO Q18]

16. Where do you work most often?

- 1 At home [SKIP to Q18]]
- 2 In an office outside your home
- 3 In some other place outside your home

[ASK ONLY IF 15a = yes or 15b = yes]

17. How did you usually get to work last week? If you use more than one method of transportation during the trip, choose the one you used for the most distance.

- 1 Personal vehicle
- 2 Public transportation
- 3 Other

[Experiment A: Questions with Qualifications

PROG. NOTE: the respondent should receive items 18, 19, and 20 in the same condition as they received item 1.]

[IF EXPERIMENT A = 1 ASK 18A, 19A, AND 20A in IVR]

[IF EXPERIMENT A = 2 ASK 18A, 19A, AND 20A in Web]

18A. Now, a question about times when you may have been away from home for personal reasons.

In the past year, how many times, if any, were you away from home on a trip? For the purposes of this question, only report times when you were away from home for 3 or more nights in a row. Exclude all business trips. If you were not away from home (as we have defined it) in the past year, enter "0."

Number of trips

19A. Here is another question about communications you may have had

In the past week, how many email messages, if any, have you written ~~or received~~? For the purposes of this question, include email messages you wrote ~~or received~~ for personal reasons. Do not include email messages you wrote ~~or received~~ for work-related reasons. If you did not write ~~or receive~~ any email messages (as we have defined them) in the past week, enter "0".

Number of emails

Redline_Channel Experiment_Questionnaire

20A. The next question is about communications you may have had with a doctor.

In the past year, how many times, if any, have you seen or talked to a medical doctor? For the purposes of this question, include the number of times you saw or talked to a general practitioner, such as a doctor in family or internal medicine. Do not include the number of times you saw or talked to a specialist, such as an obstetrician, gynecologist, or ophthalmologist. If you did not see or talk to a doctor (as we have defined it) in the past year, enter "0."

Number of times

[IF EXPERIMENT A = 3 ASK 18B, 19B and 20B IN IVR]
[IF EXPERIMENT A = 4 ASK 18B, 19B and 20B IN WEB]

18B. Now, a question about times when you may have been away from home for personal reasons.

For the purposes of this question, only report times when you were away from home for 3 or more nights in a row. Exclude all business trips. If you were not away from home (as we have defined it) in the past year, enter "0." In the past year, how many times, if any, were you away from home on a trip?

Number of trips

19B. Here is another question about communications you may have had.

For the purposes of this question, include email messages you wrote ~~or received~~ for personal reasons. Do not include email messages you wrote ~~or received~~ for work-related reasons. If you did not write ~~or receive~~ any email messages (as we have defined them) in the past week, enter "0". In the past week, how many email messages, if any, have you written ~~or received~~?

Number of emails

Redline_Channel Experiment_Questionnaire

20B. The next question is about communications you may have had with a doctor.

For the purposes of this question, include the number of times you saw or talked to a general practitioner, such as a doctor in family or internal medicine. Do not include the number of times you saw or talked to a specialist, such as an obstetrician, gynecologist, or ophthalmologist. If you did not see or talk to a doctor (as we have defined it) in the past year, enter "0." In the past year, how many times, if any, have you seen or talked to a medical doctor?

Number of times

[IF EXPERIMENT A = 5 ASK 18C,19C, AND 20C IN IVR]

[IF EXPERIMENT A = 6 ASK 18C,19C, AND 20C IN WEB]

18C1. Now, a question about times when you may have been away from home for personal reasons.

In the past year, how many times, if any, were you away from home on a trip?

Number of trips

[If 18C1>0 ask 18C2. Keep 18C1 on screen when ask 18C2.]

18C2 When you reported the number of times you were away from home, how many of them, if any, were for 3 or more nights?

Number of trips

[ASK IF 18C2>0]

[Keep 18C2 on screen when ask 18C3]

18C3 When you reported the number of times you were away from home for 3 or more nights, how many of them, if any, were for business?

Number of trips

Redline_Channel Experiment_Questionnaire

19C1. Here is another question about communications you may have had.

In the past week, how many email messages, if any, have you written ~~or received~~?

Number of emails

[If 19C1>0 then ask 19C2. Keep 19C1 on screen when ask 19C2.]

19C2. When you reported the number of email messages you wrote ~~or received~~ last week, how many of them, if any, were for personal reasons?

Number of emails

[If 19C1>0 ask 19C3, unless 19C1=19C2. Keep 19C1 and 19C2 on screen when ask 19C3.]

19C3. When you reported the number of email messages you wrote ~~or received~~ last week, how many of them, if any, were for work-related reasons?

Number of emails

20C1. The next question is about communications you may have had with a doctor.

In the past year, how many times, if any, have you seen or talked to a medical doctor?

Number of times

[If 20C1>0 then ask 20C2; Keep 20C1 on screen when ask 20C2]

20C2. When you reported seeing or talking to a doctor, how many of them, if any, included seeing or talking to a general practitioner, such as a doctor in family or internal medicine?

Number of times

Redline_Channel Experiment_Questionnaire

[If 20C1>0 then ask 20C3, unless 20C1=20C2 then do not ask 20C3; Keep 20C1 and 20C2 on screen when ask 20C3]

20C3. When you reported seeing or talking to a doctor, how many of them, if any, included seeing or talking to a specialist, such as an obstetrician, gynecologist, or ophthalmologist?

Number of times

[Buffer Question: 21.22.and 23]

The next few questions are about your community.

21 How would you rate the quality of life in your community?

- 1 Excellent
- 2 Good
- 3 Fair
- 4 Poor

22. How would you rate economic conditions in your community?

- 1 Excellent
- 2 Good
- 3 Fair
- 4 Poor

23. In the past year, have you worked with others or joined an organization in your community to do something about some community problem?

- 1 Yes
- 2 No

—

[Experiment A: Questions with Qualifications

PROG. NOTE: the respondent should receive items 24 and 25 in the same condition as they received item 1.

[IF EXPERIMENT A = 1 ASK 24A and 25A IN IVR]

[IF EXPERIMENT A = 2 ASK 24A and 25A IN WEB]

24A The next question is about the bedrooms in your house, apartment, or mobile home.

How many bedrooms are in your house, apartment, or mobile home? For the purposes of this question, only include those bedrooms located on the main floor of this house, apartment, or mobile home. If you do not have any bedrooms (as we have defined them), enter '0.' If this is an efficiency/studio apartment, enter "0".

Number of bedrooms

25A. The next question is about the other rooms in your house, apartment, or mobile home.

How many other separate rooms are in your house, apartment, or mobile home? For the purposes of this question, rooms must be separated by built-in archways or walls that extend out at least 6 inches and go from floor to ceiling. Exclude bedrooms, bathrooms, laundry rooms, porches, balconies, foyers, halls, or unfinished basements.

Number of rooms

[IF EXPERIMENT A = 3 ASK 24B and 25B]

[IF EXPERIMENT A = 4 ASK 24B and 25B]

24B. The next question is about the bedrooms in your house, apartment, or mobile home.

For the purposes of this question, only include those bedrooms located on the main floor of your house, apartment, or mobile home. If you do not have any bedrooms (as we have defined them), enter '0.' If this is an efficiency/studio apartment, enter "0."
How many bedrooms are in this house, apartment, or mobile home?

Redline_Channel Experiment_Questionnaire

Number of bedrooms

25B. The next question is about the other rooms in your house, apartment, or mobile home.

For the purposes of this question, rooms must be separated by built-in archways or walls that extend out at least 6 inches and go from floor to ceiling. Exclude bedrooms, bathrooms, laundry rooms, porches, balconies, foyers, halls, or unfinished basements. How many other separate rooms are in your house, apartment, or mobile home?

Number of rooms

[IF EXERIMENT A = 5 ASK 24C AND 25C IN IVR]
[IF EXERIMENT A = 6 ASK 24C AND 25C IN WEB]

24C1. The next question is about bedrooms in your house, apartment, or mobile home.

How many bedrooms are in your house, apartment, or mobile home?

Number of bedrooms

[ASK ONLY IF 24C1>0]

[Keep 24C1 on the screen when ask 24C2]

24C2 When you reported the number of bedrooms in your house, apartment or mobile home, how many bedrooms located on the main floor, if any, were included?

Number of bedrooms

[If 24C1=0 or 24C1=1, then ask 24C3]

24C3. Is this is an efficiency or studio apartment?

- 1 Yes
- 2 No

25C1 The next question is about the other rooms in your house, apartment, or mobile home.

How many other rooms are in your house, apartment, or mobile home that are separated by built-in archways or have walls that extend out at least 6 inches and go from floor to ceiling?

Number of rooms

[If 25C1>0 ask 25C2. Keep 25C1 on the screen when ask 25C2]

25C2 When you reported the number of other rooms in your house, apartment, or mobile home, how many bedrooms, bathrooms, laundry rooms, porches, balconies, foyers, halls, or unfinished basements, if any, were included?

Number of rooms

26. One goal of this survey is to find ways to improve future surveys. This last question is very important to our understanding of your experience during the completion of this survey.

For IVR:

On a scale of 1 to 5, where 1 means 'very slow' and 5 means 'very fast,' how would you rate the speed of this survey?

For Web:

On a scale of 1 to 5, where 1 means 'very slow' and 5 means 'very fast,' how would you rate the speed at which you took this survey?

1 Very slow

2

3

4

5 Very fast

ASK Q.M2-M3 FOR IVR VERSION ONLY

M2. May I please have your first and last name. This information will not be used to identify you. Please say and spell your first name and last name out loud and press the pound key after you have finished speaking your response.

M3. To complete the survey, please enter on your keypad the ten digit phone number including the area code that we reached you again.

ASK FOR WEB VERSION ONLY

Thank you for completing the survey. We'd like to send you \$10 for participating. So we can email you an electronic debit card, can you please provide your email address? You should receive your \$10 the next business day.

EMAIL:

I decline/I prefer not to provide email address

[ENDING MAIN SURVEY TIME STAMP]

XX. CATI: INTERVIEW CLASSIFICATION

- 1> **Completed WEB interview (auto punch at the end of WEB interview)**
- 2> **Completed IVR interview (auto punch at the end of IVR interview)**
- 3> Web Users: Web Screener Complete/Provided Email Info (W3=1-2)
- 4> Web Users Refused to Participate in Web Survey (W1=2)
- 5> Web Users: Refused email info for Web Survey (W5=3)
- 6> Web Users: Does not have email/Refused to Participate in Web Survey (W2=2 AND W5=3)
- 7> Web Users: Requested fax for link and ID (W5=1)
- 8> Web Users: Read link & ID (W5=2)
- 9> Completed IVR screener: Transferred to IVR (I2=1)
- 10> Completed IVR screener: Refused IVR interview (IF I1=2 or I2=2)

CATI: EVERYBODY AT Q.XX COUNT AS COMPLETED INTERVIEW

[TIME STAMP]

END - Thank you very much for participating in this survey.

A.3 Experiment 3

DESIGN: SET VARIABLE
PROG. NOTE: RANDOMLY ASSIGN CELL
PROG. NOTE: RE-ENABLE THE "PREVIOUS" BUTTON FOR THIS AND SUBSEQUENT
EXPERIMENTS

ROGER3A. What experimental treatment will R get?

- 1 High Frequency, Broad Category, Before Question, Same Font, Block Text
- 2 High Frequency, Narrow Category, Before Question, Same Font, Block Text
- 3 High Frequency, Broad Category, Before Question, Different Font, Block Text
- 4 High Frequency, Narrow Category, Before Question, Different Font, Block Text
- 5 High Frequency, Broad Category, After Question, Same Font, Block Text
- 6 High Frequency, Narrow Category, After Question, Same Font, Block Text
- 7 High Frequency, Broad Category, After Question, Different Font, Block Text
- 8 High Frequency, Narrow Category, After Question, Different Font, Block Text
- 9 Low Frequency, Broad Category, Before Question, Same Font, Block Text
- 10 Low Frequency, Narrow Category, Before Question, Same Font, Block Text
- 11 Low Frequency, Broad Category, Before Question, Different Font, Block Text
- 12 Low Frequency, Narrow Category, Before Question, Different Font, Block Text
- 13 Low Frequency, Broad Category, After Question, Same Font, Block Text
- 14 Low Frequency, Narrow Category, After Question, Same Font, Block Text
- 15 Low Frequency, Broad Category, After Question, Different Font, Block Text
- 16 Low Frequency, Narrow Category, After Question, Different Font, Block Text
- 17 High Frequency, Broad Category, Before Question, Same Font, Bulleted List
- 18 High Frequency, Narrow Category, Before Question, Same Font, Bulleted List
- 19 High Frequency, Broad Category, Before Question, Different Font, Bulleted List
- 20 High Frequency, Narrow Category, Before Question, Different Font, Bulleted List
- 21 High Frequency, Broad Category, After Question, Same Font, Bulleted List
- 22 High Frequency, Narrow Category, After Question, Same Font, Bulleted List
- 23 High Frequency, Broad Category, After Question, Different Font, Bulleted List
- 24 High Frequency, Narrow Category, After Question, Different Font, Bulleted List
- 25 Low Frequency, Broad Category, Before Question, Same Font, Bulleted List
- 26 Low Frequency, Narrow Category, Before Question, Same Font, Bulleted List
- 27 Low Frequency, Broad Category, Before Question, Different Font, Bulleted List
- 28 Low Frequency, Narrow Category, Before Question, Different Font, Bulleted List
- 29 Low Frequency, Broad Category, After Question, Same Font, Bulleted List
- 30 Low Frequency, Narrow Category, After Question, Same Font, Bulleted List
- 31 Low Frequency, Broad Category, After Question, Different Font, Bulleted List
- 32 Low Frequency, Narrow Category, After Question, Different Font, Bulleted List

break

Next are some questions about foods you may usually eat or drink.

break

{IF ROGER3A=1, ASK Q23A; OTHERWISE GO TO FILTER BEFORE Q23B}
DESIGN: RANDOMIZE QUESTIONS Q23A–Q26A

Q23A. For the purposes of this question, meat includes beef, pork, poultry, and other meat. How many servings of meat do you typically eat each week?

servings: [RECORD NUMBER 0–999]

break

Q24A. For the purposes of this question, dairy products include milk, cheese, yogurt, and other dairy products. How many servings of dairy products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q25A. For the purposes of this question, grain products include bread, pasta, rice, and other grain products. How many servings of grain products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q26A. For the purposes of this question, fruits and vegetables include apples, bananas, lettuce and other fruits and vegetables. How many servings of fruits and vegetables do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=2, ASK Q23B; OTHERWISE GO TO FILTER BEFORE Q23C}
DESIGN: RANDOMIZE QUESTIONS Q23B–Q26B

Q23B. For the purposes of this question, poultry includes chicken, turkey, duck, and other poultry. How many servings of poultry do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q24B. For the purposes of this question, cheese includes cheddar cheese, Swiss cheese, cottage cheese, and other cheese. How many servings of cheese do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q25B. For the purposes of this question, bread includes white bread, French bread, corn bread and other bread. How many servings of bread do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q26B. For the purposes of this question, vegetables include lettuce, tomatoes, carrots, and other vegetables. How many servings of vegetables do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=3, ASK Q23C; OTHERWISE GO TO FILTER BEFORE Q23D}
DESIGN: RANDOMIZE QUESTIONS Q23C–Q26C

Q23C. *For the purposes of this question, meat includes beef, pork, poultry, and other meat.* How many servings of meat do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q24C. *For the purposes of this question, dairy products include milk, cheese, yogurt, and other dairy products.* How many servings of dairy products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q25C. *For the purposes of this question, grain products include bread, pasta, rice, and other grain products.* How many servings of grain products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q26C. *For the purposes of this question, fruits and vegetables include apples, bananas, lettuce and other fruits and vegetables.* How many servings of fruits and vegetables do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=4, ASK Q23D; OTHERWISE GO TO FILTER BEFORE Q23E}
DESIGN: RANDOMIZE QUESTIONS Q23D–Q26D

Q23D. *For the purposes of this question, poultry includes chicken, turkey, duck, and other poultry.* How many servings of poultry do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q24D. *For the purposes of this question, cheese includes cheddar cheese, Swiss cheese, cottage cheese, and other cheese.* How many servings of cheese do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q25D. *For the purposes of this question, bread includes white bread, French bread, corn bread and other bread.* How many servings of bread do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q26D. *For the purposes of this question, vegetables include lettuce, tomatoes, carrots, and other vegetables.* How many servings of vegetables do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break

{IF ROGER3A=5, ASK Q23E; OTHERWISE GO TO FILTER BEFORE Q23F}
DESIGN: RANDOMIZE QUESTIONS Q23E–Q26E

Q23E. How many servings of meat do you typically eat each week? For the purposes of this question, meat includes beef, pork, poultry, and other meat.

servings: [RECORD NUMBER 0–999]
break

Q24E. How many servings of dairy products do you typically eat each week? For the purposes of this question, dairy products include milk, cheese, yogurt, and other dairy products.

servings: [RECORD NUMBER 0–999]
break

Q25E. How many servings of grain products do you typically eat each week? For the purposes of this question, grain products include bread, pasta, rice, and other grain products.

servings: [RECORD NUMBER 0–999]
break

Q26E. How many servings of fruits and vegetables do you typically eat each week? For the purposes of this question, fruits and vegetables include apples, bananas, lettuce and other fruits and vegetables.

servings: [RECORD NUMBER 0–999]
break

{IF ROGER3A=6, ASK Q23F; OTHERWISE GO TO FILTER BEFORE Q23G}
DESIGN: RANDOMIZE QUESTIONS Q23F–Q26F

Q23F. How many servings of poultry do you typically eat each week? For the purposes of this question, poultry includes chicken, turkey, duck, and other poultry.

servings: [RECORD NUMBER 0–999]
break

Q24F. How many servings of cheese do you typically eat each week? For the purposes of this question, cheese includes cheddar cheese, Swiss cheese, cottage cheese, and other cheese.

servings: [RECORD NUMBER 0–999]
break

Q25F. How many servings of bread do you typically eat each week? For the purposes of this question, bread includes white bread, French bread, corn bread and other bread.

servings: [RECORD NUMBER 0–999]
break

Q26F. How many servings of vegetables do you typically eat each week? For the purposes of this question, vegetables include lettuce, tomatoes, carrots, and other vegetables.

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=7, ASK Q23G; OTHERWISE GO TO FILTER BEFORE Q23H}
DESIGN: RANDOMIZE QUESTIONS Q23G–Q26G

Q23G. How many servings of meat do you typically eat each week? *For the purposes of this question, meat includes beef, pork, poultry, and other meat.*

servings: [RECORD NUMBER 0–999]
break _____

Q24G. How many servings of dairy products do you typically eat each week? *For the purposes of this question, dairy products include milk, cheese, yogurt, and other dairy products.*

servings: [RECORD NUMBER 0–999]
break _____

Q25G. How many servings of grain products do you typically eat each week? *For the purposes of this question, grain products include bread, pasta, rice, and other grain products.*

servings: [RECORD NUMBER 0–999]
break _____

Q26G. How many servings of fruits and vegetables do you typically eat each week? *For the purposes of this question, fruits and vegetables include apples, bananas, lettuce and other fruits and vegetables.*

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=8, ASK Q23H; OTHERWISE GO TO FILTER BEFORE Q23I}
DESIGN: RANDOMIZE QUESTIONS Q23H–Q26H

Q23H. How many servings of poultry do you typically eat each week? *For the purposes of this question, poultry includes chicken, turkey, duck, and other poultry.*

servings: [RECORD NUMBER 0–999]
break _____

Q24H. How many servings of cheese do you typically eat each week? *For the purposes of this question, cheese includes cheddar cheese, Swiss cheese, cottage cheese, and other cheese.*

servings: [RECORD NUMBER 0–999]
break _____

Q25H. How many servings of bread do you typically eat each week? *For the purposes of this question, bread includes white bread, French bread, corn bread and other bread.*

servings: [RECORD NUMBER 0–999]
break _____

Q26H. How many servings of vegetables do you typically eat each week? *For the purposes of this question, vegetables include lettuce, tomatoes, carrots, and other vegetables.*

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=9, ASK Q23I; OTHERWISE GO TO FILTER BEFORE Q23J}
DESIGN: RANDOMIZE QUESTIONS Q23I–Q26I

Q23I. For the purposes of this question, meat includes lamb, veal, goat, and other meat. How many servings of meat do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q24I. For the purposes of this question, dairy products include frozen yogurt, feta cheese, custard, and other dairy products. How many servings of dairy products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q25I. For the purposes of this question, grain products include millet, puffed wheat, couscous, and other grain products. How many servings of grain products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q26I. For the purposes of this question, fruits and vegetables include grapefruit, dried fruit, asparagus, and other fruits and vegetables. How many servings of fruits and vegetables do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=10, ASK Q23J; OTHERWISE GO TO FILTER BEFORE Q23K}
DESIGN: RANDOMIZE QUESTIONS Q23J–Q26J

Q23J. For the purposes of this question, poultry includes goose, quail, pheasant, and other poultry. How many servings of poultry do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q24J. For the purposes of this question, cheese includes blue cheese, Brie, Gouda, and other cheese. How many servings of cheese do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q25J. For the purposes of this question, bread includes wheat bread, whole grain bread, brioche, and other bread. How many servings of bread do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q26J. For the purposes of this question, vegetables include asparagus, Brussel sprouts, green beans and other vegetables. How many servings of vegetables do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=11, ASK Q23K; OTHERWISE GO TO FILTER BEFORE Q23L}
DESIGN: RANDOMIZE QUESTIONS Q23K–Q26K

Q23K. *For the purposes of this question, meat includes lamb, veal, goat, and other meat.* How many servings of meat do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q24K. *For the purposes of this question, dairy products include frozen yogurt, feta cheese, custard, and other dairy products.* How many servings of dairy products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q25K. *For the purposes of this question, grain products include millet, puffed wheat, couscous, and other grain products.* How many servings of grain products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q26K. *For the purposes of this question, fruits and vegetables include grapefruit, dried fruit, asparagus, and other fruits and vegetables.* How many servings of fruits and vegetables do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=12, ASK Q23L; OTHERWISE GO TO FILTER BEFORE Q23M}
DESIGN: RANDOMIZE QUESTIONS Q23L–Q26L

- Q23L. *For the purposes of this question, poultry includes goose, quail, pheasant, and other poultry.* How many servings of poultry do you typically eat each week?
- # servings: [RECORD NUMBER 0–999]
break _____
-
- Q24L. *For the purposes of this question, cheese includes blue cheese, Brie, Gouda, and other cheese.* How many servings of cheese do you typically eat each week?
- # servings: [RECORD NUMBER 0–999]
break _____
-
- Q25L. *For the purposes of this question, bread includes wheat bread, whole grain bread, brioche, and other bread.* How many servings of bread do you typically eat each week?
- # servings: [RECORD NUMBER 0–999]
break _____
-
- Q26L. *For the purposes of this question, vegetables include asparagus, Brussel sprouts, green beans and other vegetables.* How many servings of vegetables do you typically eat each week?
- # servings: [RECORD NUMBER 0–999]
break _____
-

{IF ROGER3A=13, ASK Q23M; OTHERWISE GO TO FILTER BEFORE Q23N}
DESIGN: RANDOMIZE QUESTIONS Q23M–Q26M

- Q23M. How many servings of meat do you typically eat each week? For the purposes of this question, meat includes lamb, veal, goat, and other meat.
- # servings: [RECORD NUMBER 0–999]
break _____
-
- Q24M. How many servings of dairy products do you typically eat each week? For the purposes of this question, dairy products include frozen yogurt, feta cheese, custard, and other dairy products.
- # servings: [RECORD NUMBER 0–999]
break _____
-
- Q25M. How many servings of grain products do you typically eat each week? For the purposes of this question, grain products include millet, puffed wheat, couscous, and other grain products.
- # servings: [RECORD NUMBER 0–999]
break _____
-
- Q26M. How many servings of fruits and vegetables do you typically eat each week? For the purposes of this question, fruits and vegetables include grapefruit, dried fruit, asparagus, and other fruits and vegetables.
- # servings: [RECORD NUMBER 0–999]

break

{IF ROGER3A=14, ASK Q23N; OTHERWISE GO TO FILTER BEFORE Q23O}
DESIGN: RANDOMIZE QUESTIONS Q23N–Q26N

Q23N. How many servings of poultry do you typically eat each week? For the purposes of this question, poultry includes goose, quail, pheasant, and other poultry.

servings: [RECORD NUMBER 0–999]

break

Q24N. How many servings of cheese do you typically eat each week? For the purposes of this question, cheese includes blue cheese, Brie, Gouda, and other cheese.

servings: [RECORD NUMBER 0–999]

break

Q25N. How many servings of bread do you typically eat each week? For the purposes of this question, bread includes wheat bread, whole grain bread, brioche, and other bread.

servings: [RECORD NUMBER 0–999]

break

Q26N. How many servings of vegetables do you typically eat each week? For the purposes of this question, vegetables include asparagus, Brussel sprouts, green beans and other vegetables.

servings: [RECORD NUMBER 0–999]

break

{IF ROGER3A=15, ASK Q23O; OTHERWISE GO TO FILTER BEFORE Q23P}
DESIGN: RANDOMIZE QUESTIONS Q23O–Q26O

Q23O. How many servings of meat do you typically eat each week? *For the purposes of this question, meat includes lamb, veal, goat, and other meat.*

servings: [RECORD NUMBER 0–999]

break

Q24O. How many servings of dairy products do you typically eat each week? *For the purposes of this question, dairy products include frozen yogurt, feta cheese, custard, and other dairy products.*

servings: [RECORD NUMBER 0–999]

break

Q25O. How many servings of grain products do you typically eat each week? *For the purposes of this question, grain products include millet, puffed wheat, couscous, and other grain products.*

servings: [RECORD NUMBER 0–999]

break

Q26O. How many servings of fruits and vegetables do you typically eat each week? *For the purposes of this question, fruits and vegetables include grapefruit, dried fruit, asparagus, and other fruits and vegetables.*

servings: [RECORD NUMBER 0–999]
break

{IF ROGER3A=16, ASK Q23P; OTHERWISE GO TO FILTER BEFORE Q23Q}
DESIGN: RANDOMIZE QUESTIONS Q23P–Q26P

Q23P. How many servings of poultry do you typically eat each week? *For the purposes of this question, poultry includes goose, quail, pheasant, and other poultry.*

servings: [RECORD NUMBER 0–999]
break

Q24P. How many servings of cheese do you typically eat each week? *For the purposes of this question, cheese includes blue cheese, Brie, Gouda, and other cheese.*

servings: [RECORD NUMBER 0–999]
break

Q25P. How many servings of bread do you typically eat each week? *For the purposes of this question, bread includes wheat bread, whole grain bread, brioche, and other bread.*

servings: [RECORD NUMBER 0–999]
break

Q26P. How many servings of vegetables do you typically eat each week? *For the purposes of this question, vegetables include asparagus, Brussel sprouts, green beans and other vegetables.*

servings: [RECORD NUMBER 0–999]
break

{IF ROGER3A=17, ASK Q23Q; OTHERWISE GO TO FILTER BEFORE Q23R}
DESIGN: RANDOMIZE QUESTIONS Q23Q–Q26Q

Q23Q. For the purposes of this question, meat includes:
beef,
pork,
poultry,
other meat.

How many servings of meat do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break

Q24Q. For the purposes of this question, dairy products include:
milk,
cheese,
yogurt,
other dairy products.

How many servings of dairy products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q25Q. For the purposes of this question, grain products include:
bread,
pasta,
rice,
other grain products.
How many servings of grain products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q26Q. For the purposes of this question, fruits and vegetables include:
apples,
bananas,
lettuce,
other fruits and vegetables.
How many servings of fruits and vegetables do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=18, ASK Q23R; OTHERWISE GO TO FILTER BEFORE Q23S}
DESIGN: RANDOMIZE QUESTIONS Q23R–Q26R

Q23R. For the purposes of this question, poultry includes:
chicken,
turkey,
duck,
other poultry.
How many servings of poultry do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q24R. For the purposes of this question, cheese includes:
cheddar cheese,
Swiss cheese,
cottage cheese,
other cheese.
How many servings of cheese do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q25R. For the purposes of this question, bread includes:
white bread,
French bread,
corn bread,
other bread.
How many servings of bread do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q26R. For the purposes of this question, vegetables include:
lettuce,
tomatoes,
carrots,
other vegetables.
How many servings of vegetables do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=19, ASK Q23S; OTHERWISE GO TO FILTER BEFORE Q23T}
DESIGN: RANDOMIZE QUESTIONS Q23S–Q26S

Q23S. *For the purposes of this question, meat includes:*
beef,
pork,
poultry,
other meat.
How many servings of meat do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q24S. *For the purposes of this question, dairy products include:*
milk,
cheese,
yogurt,
other dairy products.
How many servings of dairy products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q25S. *For the purposes of this question, grain products include:*
bread,
pasta,
rice,
other grain products.
How many servings of grain products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q26S. *For the purposes of this question, fruits and vegetables include:*
apples,
bananas,
lettuce,
other fruits and vegetables.
How many servings of fruits and vegetables do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

break

Q24U. How many servings of dairy products do you typically eat each week? For the purposes of this question, dairy products include:
milk,
cheese,
yogurt,
other dairy products.

servings: [RECORD NUMBER 0–999]
break

Q25U. How many servings of grain products do you typically eat each week? For the purposes of this question, grain products include:
bread,
pasta,
rice,
other grain products.

servings: [RECORD NUMBER 0–999]
break

Q26U. How many servings of fruits and vegetables do you typically eat each week? For the purposes of this question, fruits and vegetables include:
apples,
bananas,
lettuce,
other fruits and vegetables.

servings: [RECORD NUMBER 0–999]
break

{IF ROGER3A=22, ASK Q23V; OTHERWISE GO TO FILTER BEFORE Q23W}
DESIGN: RANDOMIZE QUESTIONS Q23V–Q26V

Q23V. How many servings of poultry do you typically eat each week? For the purposes of this question, poultry includes:
chicken,
turkey,
duck,
other poultry.

servings: [RECORD NUMBER 0–999]
break

Q24V. How many servings of cheese do you typically eat each week? For the purposes of this question, cheese includes:
cheddar cheese,
Swiss cheese,
cottage cheese,
other cheese.

servings: [RECORD NUMBER 0–999]
break

Q25V. How many servings of bread do you typically eat each week? For the purposes of this question, bread includes:
white bread,
French bread,
corn bread,
other bread.

servings: [RECORD NUMBER 0–999]
break _____

Q26V. How many servings of vegetables do you typically eat each week? For the purposes of this question, vegetables include:
lettuce,
tomatoes,
carrots,
other vegetables.

servings: [RECORD NUMBER 0–999]
break _____

{IF roger3a=23, ASK Q23W; OTHERWISE GO TO FILTER BEFORE Q23X}
DESIGN: RANDOMIZE QUESTIONS Q23W–Q26W

Q23W. How many servings of meat do you typically eat each week? *For the purposes of this question, meat includes:*
beef,
pork,
poultry,
other meat.

servings: [RECORD NUMBER 0–999]
break _____

Q24W. How many servings of dairy products do you typically eat each week? *For the purposes of this question, dairy products include:*
milk,
cheese,
yogurt,
other dairy products.

servings: [RECORD NUMBER 0–999]
break _____

Q25W. How many servings of grain products do you typically eat each week? *For the purposes of this question, grain products include:*
bread,
pasta,
rice,
other grain products.

servings: [RECORD NUMBER 0–999]
break _____

Q26W. How many servings of fruits and vegetables do you typically eat each week? *For the purposes of this question, fruits and vegetables include:*
apples,
bananas,
lettuce,
other fruits and vegetables.

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=24, ASK Q23X; OTHERWISE GO TO FILTER BEFORE Q23Y}
DESIGN: RANDOMIZE QUESTIONS Q23X–Q26X

Q23X. How many servings of poultry do you typically eat each week? *For the purposes of this question, poultry includes:*
chicken,
turkey,
duck,
other poultry.

servings: [RECORD NUMBER 0–999]
break _____

Q24X. How many servings of cheese do you typically eat each week? *For the purposes of this question, cheese includes:*
cheddar cheese,
Swiss cheese,
cottage cheese,
other cheese.

servings: [RECORD NUMBER 0–999]
break _____

Q25X. How many servings of bread do you typically eat each week? *For the purposes of this question, bread includes:*
white bread,
French bread,
corn bread,
other bread.

servings: [RECORD NUMBER 0–999]
break _____

Q26X. How many servings of vegetables do you typically eat each week? *For the purposes of this question, vegetables include:*
lettuce,
tomatoes,
carrots,
other vegetables.

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=25, ASK Q23Y; OTHERWISE GO TO FILTER BEFORE Q23Z}
DESIGN: RANDOMIZE QUESTIONS Q23Y–Q26Y

Q23Y. For the purposes of this question, meat includes:
lamb,
veal,
goat,
other meat.
How many servings of meat do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q24Y. For the purposes of this question, dairy products include:
frozen yogurt,
feta cheese,
custard,
other dairy products.
How many servings of dairy products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q25Y. For the purposes of this question, grain products include:
millet,
puffed wheat,
couscous,
other grain products.
How many servings of grain products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q26Y. For the purposes of this question, fruits and vegetables include:
grapefruit,
dried fruit,
asparagus,
other fruits and vegetables.
How many servings of fruits and vegetables do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=26, ASK Q23Z; OTHERWISE GO TO FILTER BEFORE Q23AA}
DESIGN: RANDOMIZE QUESTIONS Q23Z–Q26Z

Q23Z. For the purposes of this question, poultry includes:
goose,
quail,
pheasant,
other poultry.
How many servings of poultry do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q24Z. For the purposes of this question, cheese includes:
blue cheese,
Brie,
Gouda,
other cheese.

How many servings of cheese do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q25Z. For the purposes of this question, bread includes:
wheat bread,
whole grain bread,
brioche,
other bread.

How many servings of bread do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q26Z. For the purposes of this question, vegetables include:
asparagus,
Brussel sprouts,
green beans,
other vegetables.

How many servings of vegetables do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=27, ASK Q23AA; OTHERWISE GO TO FILTER BEFORE Q23BB}
DESIGN: RANDOMIZE QUESTIONS Q23AA–Q26AA

Q23AA. For the purposes of this question, meat includes:
lamb,
veal,
goat,
other meat.

How many servings of meat do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q24AA. For the purposes of this question, dairy products include:
frozen yogurt,
feta cheese,
custard,
other dairy products.

How many servings of dairy products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q25AA. For the purposes of this question, grain products include:

*millet,
puffed wheat,
couscous,
other grain products.*

How many servings of grain products do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q26AA. *For the purposes of this question, fruits and vegetables include:
grapefruit,
dried fruit,
asparagus,
other fruits and vegetables.*

How many servings of fruits and vegetables do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=28, ASK Q23BB; OTHERWISE GO TO FILTER BEFORE Q23CC}
DESIGN: RANDOMIZE QUESTIONS Q23BB–Q26BB

Q23BB. *For the purposes of this question, poultry includes:
goose,
quail,
pheasant,
other poultry.*

How many servings of poultry do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q24BB *For the purposes of this question, cheese includes:
blue cheese,
Brie,
Gouda,
other cheese.*

How many servings of cheese do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q25BB *For the purposes of this question, bread includes:
wheat bread,
whole grain bread,
brioche,
other bread.*

How many servings of bread do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break _____

Q26BB. *For the purposes of this question, vegetables include:
asparagus,
Brussel sprouts,*

*green beans,
other vegetables.*

How many servings of vegetables do you typically eat each week?

servings: [RECORD NUMBER 0–999]
break

{IF ROGER3A=29, ASK Q23CC; OTHERWISE GO TO FILTER BEFORE Q23DD}
DESIGN: RANDOMIZE QUESTIONS Q23CC–Q26CC

Q23CC. How many servings of meat do you typically eat each week? For the purposes of this question, meat includes:

lamb,
veal,
goat,
other meat.

servings: [RECORD NUMBER 0–999]
break

Q24CC. How many servings of dairy products do you typically eat each week? For the purposes of this question, dairy products include:

frozen yogurt,
feta cheese,
custard,
other dairy products.

servings: [RECORD NUMBER 0–999]
break

Q25CC. How many servings of grain products do you typically eat each week? For the purposes of this question, grain products include:

millet,
puffed wheat,
couscous,
other grain products.

servings: [RECORD NUMBER 0–999]
break

Q26CC. How many servings of fruits and vegetables do you typically eat each week? For the purposes of this question, fruits and vegetables include:

grapefruit,
dried fruit,
asparagus,
other fruits and vegetables.

servings: [RECORD NUMBER 0–999]
break

{IF ROGER3A=30, ASK Q23DD; OTHERWISE GO TO FILTER BEFORE Q23EE}
DESIGN: RANDOMIZE QUESTIONS Q23DD–Q26DD

Q23DD. How many servings of poultry do you typically eat each week? For the purposes of this question, poultry includes:

goose,
quail,
pheasant,
other poultry.

servings: [RECORD NUMBER 0–999]
break _____

Q24DD. How many servings of cheese do you typically eat each week? For the purposes of this question, cheese includes:
blue cheese,
Brie,
Gouda,
other cheese.

servings: [RECORD NUMBER 0–999]
break _____

Q25DD. How many servings of bread do you typically eat each week? For the purposes of this question, bread includes:
wheat bread,
whole grain bread,
brioche,
other bread.

servings: [RECORD NUMBER 0–999]
break _____

Q26DD. How many servings of vegetables do you typically eat each week? For the purposes of this question, vegetables include:
asparagus,
Brussel sprouts,
green beans,
other vegetables.

servings: [RECORD NUMBER 0–999]
break _____

{IF ROGER3A=31, ASK Q23EE; OTHERWISE GO TO FILTER BEFORE Q23FF}
DESIGN: RANDOMIZE QUESTIONS Q23EE–Q26EE

Q23EE. How many servings of meat do you typically eat each week? *For the purposes of this question, meat includes:*
lamb,
veal,
goat,
other meat.

servings: [RECORD NUMBER 0–999]
break _____

Q24EE. How many servings of dairy products do you typically eat each week? *For the purposes of this question, dairy products include:*
frozen yogurt,
feta cheese,

*custard,
other dairy products.*

servings: [RECORD NUMBER 0–999]
break

Q25EE. How many servings of grain products do you typically eat each week? *For the purposes of this question, grain products include:*
*millet,
puffed wheat,
couscous,
other grain products.*

servings: [RECORD NUMBER 0–999]
break

Q26EE. How many servings of fruits and vegetables do you typically eat each week? *For the purposes of this question, fruits and vegetables include:*
*grapefruit,
dried fruit,
asparagus,
other fruits and vegetables.*

servings: [RECORD NUMBER 0–999]
break

{IF ROGER3A=32, ASK Q23FF; OTHERWISE GO TO ROGER4A}
DESIGN: RANDOMIZE QUESTIONS Q23FF–Q26FF

Q23FF. How many servings of poultry do you typically eat each week? *For the purposes of this question, poultry includes:*
*goose,
quail,
pheasant,
other poultry.*

servings: [RECORD NUMBER 0–999]
break

Q24FF. How many servings of cheese do you typically eat each week? *For the purposes of this question, cheese includes:*
*blue cheese,
Brie,
Gouda,
other cheese.*

servings: [RECORD NUMBER 0–999]
break

Q25FF. How many servings of bread do you typically eat each week? *For the purposes of this question, bread includes:*
*wheat bread,
whole grain bread,
brioche,
other bread.*

servings: [RECORD NUMBER 0-999]
break _____

Q26FF. How many servings of vegetables do you typically eat each week? *For the purposes of this question, vegetables include:*
asparagus,
Brussel sprouts,
green beans,
other vegetables.

servings: [RECORD NUMBER 0-999]
break _____

Bibliography

- Belson, W. (1981). *The Design and Understanding of Survey Questions*. Aldershot, England: Gower.
- Biederman, I., Subramaniam, S., Bar, M., Kalocsai, P., & Fisher, J. (1999). Subordinate-level object classification reexamined. *Psychological Research, 62*, 131–153.
- Bishop, G., Hippler, H., Schwarz, N., & Strack, F. (1988). A comparison of response effects in self-administered and telephone surveys. In R. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls II, & J. Waksberg (Eds.) *Telephone Survey Methodology*, (pp. 321–340). New York: Wiley.
- Bishop, G., & Smith, A. (2001). Response-order effects and the early Gallup split ballots. *Public Opinion Quarterly, 65*, 479–505.
- Blair, E., & Burton, S. (1987). Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research, 14*, 280–288.
- Bodner-Montville, J., Ahuja, J., Ingerwersen, L., Haggery, E., Enns, C., & Perloff, B. (2006). USDA food and nutrient database for dietary studies: Released on the Web. *Journal of Food Composition and Analysis, 19*, S100–S107.
- Buck, G. (2001). *Assessing Listening*. New York: Cambridge University Press.
- Cannell, C., Miller, P., & Oksenberg, L. (1981). Research on interviewing techniques. *Sociological Methodology, 12*, 389–437.
- Chaiken, S., & Eagly, A. (1983). Communication modality as a determinant of persuasion: The role of communicator salience. *Journal of Personality and Social Psychology, 45*, 241–256.
- Chang, L., & Krosnick, J. (2009). National surveys via RDD telephone interviewing versus the Internet: Comparing sample representativeness and response quality. *Public Opinion Quarterly, 73*, 641–678.
- Christian, L., & Dillman, D. (2004). The influence of graphical and symbolic language manipulations on responses to self-administered questions. *Public Opinion Quarterly, 68*, 57–80.
- Christian, L., Dillman, D., & Smyth, J. (2007). The effects of mode and format on answers to scalar questions in telephone and web surveys. In J. Lepkowski, C. Tucker, J. Brick, E. DeLeeuw, L. Japac, P. Lavrakas, M. Link, & R. Sangster (Eds.) *Advances in Telephone Survey Methodology*, (pp. 250–275). New York: Wiley.

- Clark, H., & Brennan, S. (1991). Grounding in communication. In L. Resnick, J. Levine, & S. Teasley (Eds.) *Perspectives on Socially Shared Cognition*, (pp. 127–149). Washington DC: American Psychological Association.
- Clark, H., & Schober, M. (1992). Asking questions and influencing answers. In J. Tanur (Ed.) *Questions About Questions: Inquiries into the Cognitive Bases of Surveys*, (pp. 15–48). New York: Russel Sage.
- Cohen, P. (1984). The pragmatics of referring and the modality of communication. *Computational Linguistics*, *10*, 97–146.
- Conrad, F., & Couper, M. (2004). Usability, comparability, and data quality across modes and technologies in census data collection. Tech. rep., Prepared for the U.S. Census Bureau through the Gunnison Consulting Group and the University of Michigan.
- Conrad, F., Couper, M., Tourangeau, R., & Peytchev, A. (2006). Use and non-use of clarification features in web surveys. *Journal of Official Statistics*, *22*, 245–269.
- Conrad, F., & Schober, M. (2000). Clarifying question meaning in a household telephone survey. *Public Opinion Quarterly*, *64*, 1–28.
- Conrad, F., Schober, M., & Coiner, T. (2007). Bringing features of human dialogue to web surveys. *Applied Cognitive Psychology*, *21*, 165–188.
- Conrad, F. G., Brown, N., & Cashman, E. (1998). Strategies for estimating behavioural frequency in survey interveiwes. *Memory*, *6*, 339–366.
- Couper, M. (2008). *Designing Effective Web Surveys*. Cambridge, New York: Cambridge University.
- De Leeuw, E. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, *21*, 233–255.
- De Ruiter, J., Mitterer, H., & Enfield, N. (2006). Projecting the end of a speakers turn: A cognitive cornerstone of conversation. *Language*, *82*(3), 515–535.
- DeLeeuw, E., & Van der Zouwen, J. (1988). Data quality in telephone and face-to-face surveys: A comparative meta-analysis. In R. M. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls, & J. Waksberg (Eds.) *Telephone Survey Methodology*, (pp. 283–300). New York: Wiley.
- Dickens, M., Harwood, K., & Carter, F. A. (1955). Studies in listenability. *Speech Monograph*, *22*, 49–59.
- Dillman, D. A. (2000). *Mail and Internet Surveys. Second Edition.*, chap. Mixed Mode Surveys, (pp. 217–244). New York: Wiley.
- Foddy, W. (1996). The in-depth testing of survey questions: A critical appraisal of methods. *1996*, *30*, 361–370.

- Foddy, W. (1998). An empirical evaluation of in-depth probes used to pretest survey questions. *Sociological Methods and Research*, 27, 103–133.
- Forsyth, B., & Lessler, J. (1991). Cognitive laboratory methods: A taxonomy. In P. Biemer, R. Groves, & L. Lyberg (Eds.) *Measurement Error in Surveys*, (pp. 393–418). New York: Wiley.
- Fowler, F. (1992). How unclear terms affect survey data. *Public Opinion Quarterly*, 56, 218–231.
- Fowler, F. (1995). Improving survey questions: Design and evaluation. In *Applied Social Research Methods Series, v.38*. Thousand Oaks: Sage Publications.
- Fricker, S., Galesic, M., Tourangeau, R., & Yan, T. (2005). An experimental comparison of web and telephone surveys. *Public Opinion Quarterly*, 69, 370–392.
- Frohlich, D. (1986). On the organization of form-filling behavior. *Information Design Journal*, 5, 43–59.
- Galesic, M., Tourangeau, R., Couper, M., & Conrad, F. (2008). Eye-tracking data: New insights on response order effects and other cognitive shortcuts in survey responding. *Public Opinion Quarterly*, 72, 892–913.
- Gerber, E., Wellens, T., & Keeley, C. (1996). Who lives here? The use of vignettes in household roster research. In *Working Papers in Survey Methodology (SM 96/02)*. Washington, DC: U.S. Census Bureau.
- Graesser, A., Bommarreddy, S., Swamer, S., & Golding, J. (1996). Integrating questionnaire design with a cognitive computational model of human question answering. In N. Schwarz, & S. Sudman (Eds.) *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*, (pp. 143–174). San Francisco: Jossey-Bass Publishers.
- Graesser, A., Cai, Z., Louwrese, M., & Daniel, F. (2006). Question understanding aid (QUAID): A web facility that tests question comprehensibility. *Public Opinion Quarterly*, 70, 3–22.
- Graesser, A., Kennedy, T., Wiemer-Hastings, P., & Ottati, V. (1999). The use of computational cognitive models to improve questions on surveys and questionnaires. In M. Sirken, D. Herrman, S. Schecter, N. Schwarz, J. Tanur, & R. Tourangeau (Eds.) *Cognition and Survey Research*, (pp. 199–216). New York: Wiley.
- Gray, W., John, B., & Atwood, M. (1993). Project Ernestine: Validating a GOMS analysis for predicting and explaining real-world performance. *Human-Computer Interaction*, 8, 237–309.
- Grice, H. (1975). Logic and conversation. In P. Cole, & J. Morgan (Eds.) *Syntax and Semantics: 3 Speech Acts*, (pp. 41–58). New York: Academic Press.

- Groves, R. (1989). *Survey Errors and Survey Costs*. New York: Wiley.
- Hartley, J. (2004). Designing instructional and informational text. In D. Jonassen (Ed.) *Handbook of Research on Educational Communications and Technology*, (pp. 917–947). Mahwah, N.J: Lawrence Erlbaum Associates, 2nd ed.
- Harwood, K. (1951). An experimental comparison of listening comprehensibility with reading comprehensibility. *Speech Monographs*, *18*, 123–124.
- Heerwegh, D., & Loosveldt, G. (2008). Face-to-face versus web surveying in a high-internet-coverage population. *Public Opinion Quarterly*, *72*, 836–846.
- Holbrook, A., Cho, Y., & Johnson, T. (2006). The impact of question and respondent characteristics on comprehension and mapping difficulties. *Public Opinion Quarterly*, *70*, 565–595.
- Holbrook, A., Krosnick, J., Moore, D., & Tourangeau, R. (2007). Response order effects in dichotomous categorical questions presented orally: The impact of question and respondent attributes. *Public Opinion Quarterly*, *71*, 325–348.
- Houtkoop-Steenstra, H. (2002). Question turn format and turn-taking problems in standardized interviews. In D. Maynard, H. Houtkoop-Steenstra, N. Schaeffer, & J. van der Zouwen (Eds.) *Standardization and Tacit Knowledge: Interaction and Practice in the Survey Interview*, (pp. 243–259). New York: John Wiley and Sons, Inc.
- Jansen, C., & Steehouder, M. (1992). Forms as a source of communication problems. *Journal of Technical Writing and Communication*, *22*, 179–194.
- Jenkins, C., & Dillman, D. (1997). Towards a theory of self-administered questionnaire design. In L. Lyberg, P. Biemer, M. Collins, L. Decker, E. DeLeeuw, C. Dippo, N. Schwarz, & D. Trewin (Eds.) *Survey Measurement and Process Quality*, (pp. 165–220). New York: Wiley-Interscience.
- Johnson, K., & Mervis, C. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology*, *21*, 60–99.
- Just, M., & Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329–354.
- Kant, E. (1781). *Critique of Pure Reason Edited and Translated by P. Guyer and A. W. Wood. 1998.* Cambridge: Cambridge University Press.
- Kindermann, C., Lynch, J., & Cantor, D. (1997). Effects of the redesign on victimization estimates. Tech. rep., U.S. Department of Justice. Office of Justice Programs.
- Klein, D., & Murphy, G. (2001). The representation of polysemous words. *Journal of Memory and Language*, *45*, 259–282.

- Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: The University of Chicago Press.
- Lynch, J. (1996). The polls review: Clarifying divergent estimates of rape from two national surveys. *Public Opinion Quarterly*, *60*, 410–430.
- Martin, E. (2002). The effects of questionnaire design on reporting of detailed Hispanic origin in Census 2000 mail questionnaires. *Public Opinion Quarterly*, *66*, 583–593.
- Martin, E., Gerber, E., & Redline, C. (2004). Census 2000 testing, experimental and evaluation program synthesis report no. 17, tr-17, census 2000 alternative questionnaire experiment. Tech. rep., U.S. Census Bureau. Washington, DC 20233.
- Martin, E., Hunter Childs, J., DeMaio, T., Hill, J., Reiser, C., Gerber, E., Styles, K., & Dillman, D. (2007a). Guidelines for designing questionnaires for administration in different modes. Tech. rep., U.S. Bureau of the Census. Washington, D.C. 20233.
- Martin, E., Sheppard, D., Bentley, M., & Bennett, C. (2007b). Results of the 2003 National Census Test of race and Hispanic questions. In *Research Report Series (Survey Methodology 2007-34)*. Washington DC: U.S. Census Bureau.
- McCarthy, M., & Mothersbaugh, D. (2002). Effects of typographic factors in advertising-based persuasion: A general model and initial empirical tests. *Psychology & Marketing*, *19*, 663–691.
- Menon, G., & Yorkston, E. (2000). The use of memory and contextual cues in the formation of behavioral frequency judgments. In A. Stone, J. Turkkan, C. Bachrach, J. B. Jobe, H. Kurtzman, & V. Cain (Eds.) *The Science of Self-Report*, (pp. 63–79). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Messmer, D., & Seymour, D. (1982). The effects of branching on item nonresponse. *Public Opinion Quarterly*, *46*, 270–277.
- Murphy, G. (1997). Polysemy and the creation of novel word meanings. In T. Ward, S. Smith, & J. Vaid (Eds.) *Creative Thought: An Investigation of Conceptual Structures and Processes*, (pp. 235–265). Washington DC: American Psychological Association.
- Nunberg, G. (1979). The non-uniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy*, *3*, 143–184.
- Oksenberg, L., Cannell, C., & Kalton, G. (1991). New strategies for pretesting survey questions. *Journal of Official Statistics*, *7*, 349–365.
- Osada, N. (2004). Listening comprehension research: A brief review of the past thirty years. *Dialogue*, *3*, 53–66.

- Pinkal, M. (1995). *Logic and lexicon: the semantics of the indefinite; translated by Geoffrey Simmons*. Dordrecht; Boston: Kluwer Academic Publishers.
- Pinker, S. (2007). *The Stuff of Thought: Language as a Window into Human Nature*. New York: Viking.
- Rayner, K., & Clifton, C. (2009). Language processing in reading and speech perception is fast and incremental: Implications for event-related potential research. *Biological Psychology, 80*, 4–9.
- Redline, C., & Dillman, D. (2002). The influence of alternative visual designs on respondents' performance with branching instructions in self-administered questionnaires. In R. Groves, D. Dillman, E. Eltinge, & R. Little (Eds.) *Survey Non-response*, (pp. 179–193). New York: John Wiley and Sons, Inc.
- Redline, C., Dillman, D., Dajani, A., & Scaggs, M. (2003). Improving navigational performance in Census 2000 by altering the visually administered languages of branching instructions. *Journal of Official Statistics, 19*, 403–419.
- Roediger, H. (1974). Inhibiting effects of recall. *Memory and Cognition, 2*, 261–269.
- Rosch, E., Mervis, C., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382–439.
- Ruhl, C. (1989). *On Monosemy: A Study in Linguistic Semantics*. Albany, New York: State University of New York Press.
- Sachs, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*, 696–735.
- Sakshaug, J., Tourageau, R., Krosnick, J., Ackermann, A., Malka, A., DeBell, M., & Turakhia, C. (2009). Dispositions and outcome rates in the face-to-face internet survey platform. In *paper presented at the American Association of Public Opinion Research*.
- Scarr, H. (1993). A review of federal measurements of race and ethnicity. presented before the subcommittee on census, statistics, and postal personnel. Tech. rep., Committee on Post Office and Civil Service.
- Schaeffer, N., & Presser, S. (2003). The science of asking questions. *Annual Review of Sociology, 29*, 65–88.
- Schegloff, E., & Sacks, H. (1973). Opening up closings. *Semiotica, 8*, 289–327.
- Schober, M., & Bloom, J. (2004). Discourse cues that respondents have misunderstood survey questions. *Discourse Processes, 38*, 287–308.
- Schober, M., & Conrad, F. (1997). Does conversational interviewing reduce survey measurement error? *The Public Opinion Quarterly, 61*, 576–602.

- Schober, M., Conrad, F., & Fricker, S. (2004). Misunderstanding standardized language in research interviews. *Applied Cognitive Psychology, 18*, 169–188.
- Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods and the logic of conversation*. Hillsdale, NJ: Erlbaum.
- Schwarz, N., Grayson, C., & Knauper, B. (1998). Formal features of rating scales and the interpretation of question meaning.. *International Journal of Public Opinion Research, 10*, 177–183.
- Schwarz, N., & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. *American Journal of Evaluation, 22*, 127–160.
- Schwarz, N., Strack, F., Hippler, H., & Bishop, G. (1991a). The impact of administration mode on response effects in survey measurement. *Applied Cognitive Psychology, 5*, 193–212.
- Schwarz, N., Strack, F., & Mai, H. (1991b). Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis. *Public Opinion Research, 55*, 3–23.
- Sloman, S., Rottenstreich, Y., Wisniewski, E., Hadjichristidis, C., & Fox, C. (2004). Typical versus atypical unpacking and superadditive probability judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 573–582.
- Smith, E. E. (1995). Concepts and categorization. In E. Smith, & D. N. Osherson (Eds.) *Thinking*, (pp. 3–33). Cambridge, MA: MIT Press, 2 ed.
- Smyth, J. D., Christian, L., & Dillman, D. (2008). Does yes or no on the telephone mean the same as check-all-that-apply on the web?. *Public Opinion Quarterly, 72*, 103–113.
- Sudman, S., Bradburn, N., & Schwarz, N. (1996). *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass Publishers.
- Suessbrick, A., Schober, M., & Conrad, F. (2000). Different respondents interpret ordinary questions quite differently. In *American Statistical Association. Section on Survey Research Methods.*, (pp. 907–912). Alexandria, VA: American Statistical Association.
- Tourangeau, R. (1984). Cognitive science and survey methods: A cognitive perspective. In T. Jabine, M. Straf, J. Tanur, & R. Tourangeau (Eds.) *Cognitive Aspects of Survey Design: Building a Bridge Between Disciplines*, (pp. 73–100). Washington DC: National Academy Press.

- Tourangeau, R., Conrad, F., Arens, Z., Fricker, S., Lee, S., & Smith, E. (2006). Everyday concepts and classification errors: Judgments of disability and residence. *Journal of Official Statistics*, *22*, 385–418.
- Tourangeau, R., Conrad, F., Couper, M., Redline, C., & Ye, C. (2009). The effects of providing examples: Questions about frequencies and ethnic background. In *Paper presented at the American Association of Public Opinion Research*.
- Tourangeau, R., Conrad, F., Couper, M., & Ye, C. (2010). Providing examples in survey questions. Unpublished manuscript.
- Tourangeau, R., Couper, M. P., & Conrad, F. (2004). Spacing, position, and order: Interpretative heuristics for visual features of survey questions. *Public Opinion Quarterly*, *68*, 368–393.
- Tourangeau, R., & Rasinski, K. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin*, *103*, 299–314.
- Tourangeau, R., Rasinski, K., & Bradburn, N. (1991). Measuring happiness in surveys: A test of the subtraction hypothesis. *Public Opinion Quarterly*, *55*, 255–266.
- Tourangeau, R., Rips, L., & Rasinski, K. (2000). *The Psychology of Survey Response*. New York: Cambridge University Press.
- Tourangeau, R., & Sakshaug, R. (2010). Weighting procedures for the MRI national sample.
- Turner, C. F., Lessler, J. L., Hubbard, B., & Witt, M. (1992). Effects of mode of administration and wording on data quality. In C. Turner, J. Lesser, & J. Gfoerer (Eds.) *Survey Measurement and Drug Abuse*, (pp. 221–243). U.S. Department of Health and Human Services.
- Tversky, A., & Koehler, D. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, *101*, 547–567.
- Van der Zouwen, J., & Dijkstra, W. (2002). Testing questionnaires using interaction coding in standardized interviews. In D. Maynard, H. Houtkoop-Streenstra, N. Schaeffer, & J. van der Zouwen (Eds.) *Standardization and Tacit Knowledge*, (pp. 427–447). New York: John Wiley and Sons, Inc.
- Wei, L. (1978). The adaptive biased coin design for sequential experiments. *The Annals of Statistics*, *6*, 92–100.
- Wennerstrom, A. (2001). *The Music of Everyday Speech: Prosody and Discourse Analysis*. New York, New York: Oxford University Press.
- Wertheimer, M. (1938). *A Source Book of Gestalt Psychology*. Routledge & Kegan Paul.

Yan, T., & Tourangeau, R. (2008). Fast times and easy questions: The effects of age, experience and question complexity on web survey response times. *Applied Cognitive Psychology, 22*, 51–68.