

## CHAPTER 6

# Making Sense of Questions: An Interactional Approach

*Michael F. Schober*

*New School for Social Research*

Imagine that you are engaged in casual conversation in your living room with a few visiting relatives. Your aunt asks you “So, how many hours a week do you work?” How do you go about making sense of the question?

Obviously there are a number of cognitive and linguistic processes involved in such an apparently ordinary encounter (see Graesser and Franklin, 1990), and these processes are worth examining for those of us interested in the cognitive aspects of survey participation. Of course, survey interviewers and respondents don’t interact in the same way as your aunt interacts with you. But clarifying the similarities and differences can help us make sense of how survey respondents interpret questions.

So how do you make sense of your aunt’s question? You need to parse the continuous stream of sound in your aunt’s utterance into words, and this requires a great deal of phonological knowledge. You need to know the grammatical structure of English well enough to figure out that your aunt’s utterance had the syntactic form of a question, and you need to parse the utterance into its grammatical parts—nouns, verbs, subjects, objects, etc. You need to know the conventional meanings of the words in the utterance, like *hours*, *week*, and *work*. You need to access those meanings and combine them, using your knowledge of grammatical structure, to create a coherent sense of what the utterance means. You need to know enough about social situations to understand that when a question is asked, the questioner wants an answer.

Beyond these cognitive and linguistic processes, you also need to figure out

how the conventional meanings apply in the current context. This isn't always straightforward. You need to know whether the "you" in the question refers to yourself alone, or you and your spouse together. You need to know whether your aunt is asking for a ballpark estimate or an absolutely precise number of hours. You need to know whether your aunt wants you to think of work broadly or narrowly: Do the hours you spend socializing with the boss count as work time? If you are wary about your family, you need to know your aunt's agenda in asking the question. Is she genuinely curious or is she comparing you unfavorably to your more industrious cousin?

If you are uncertain about what she meant by the question, you are likely to ask: "Do you mean both of us?" or "Do you mean everything I do that's related to work?" or (most dangerously) "What are you getting at?" Alternatively, you can answer the question assuming one interpretation: "Well, *I* work 45 hours a week" or "*Officially* I'm paid for 40 hours of work per week" or "I just won the Employee of the Month Award." You do this on the presumption that if your aunt is not satisfied with your answer she will ask a follow-up question.

This example illustrates that making sense of questions involves more than just individual processes like computing syntactic structure or accessing conventional word meanings. Making sense of questions also involves an *interactive* element: People make inferences about the questioner's intentions by relying on assumptions about how the social world works (Grice, 1975) and they rely on the questioner to help interpret questions in subsequent dialogue. I propose that we can only understand some things people do as they make sense of questions if we consider them to be engaged in interactive processes, not just individual cognitive processes (see Clark, 1992, 1996; Schober, 1998; Schober and Conrad, in press).

## 6.1 MAKING SENSE OF QUESTIONS IN SPONTANEOUS CONVERSATION

In ordinary conversation, addressees make sense of speakers' questions by relying on at least two resources that could be called interactive. First, to arrive at an initial interpretation, addressees presume that questioners have followed a principle of *audience design*, basing the wording and framing of their questions on the knowledge, beliefs, and assumptions that both participants share. Second, addressees rely on *grounding* procedures to make sure they have understood the question. Because addressees' initial interpretations of questions aren't guaranteed to match speakers' intentions, conversational participants can engage in additional conversational turns to reach agreement that a question has been understood as intended.

### 6.1.1 Audience Design

When you hear your aunt address her question to you, you know that she has designed it with (at least) you in mind, and this affects how you interpret what

she says. For the moment you have been cast in the listener role of being her designated *addressee*, rather than in any of the other listener roles you might play in a conversation. Your relatives on the sofa who have been taking part in the conversation, but aren't currently being addressed, have been cast in the role of *side participants*. Your spouse in the kitchen, who hasn't been taking part in the conversation but can overhear it, is playing the role of *bystander*. Your neighbors who, unbeknownst to you, are listening through the paper-thin walls are taking the role of another kind of overhearer: *eavesdroppers* (see Clark, 1992, 1996; Clark and Carlson, 1982; Goffman, 1981; Schober and Clark, 1989, for discussion of listener roles).

How does your listener role affect the way you make sense of her question? As an addressee, you can assume that your aunt believes that you will be able to interpret her through the words she utters (Grice, 1957), against the background premises and knowledge that you and she share. The side participants have to make a slightly different assumption: Your aunt probably wants them to understand the question she is addressing to you, but her intentions toward them are different than her intentions toward the addressee (you). She is probably doing something like informing them of her intentions toward you (see Clark and Carlson, 1982). She could also be doing something more complicated, like trying to let the side participants know something about her attitude toward you without your knowing it. Bystanders and eavesdroppers make sense of your aunt's question on different bases still. Bystanders must understand that your aunt may not even be attempting to inform them of her intentions toward you; eavesdroppers are certain that she is not.

As you can see, inferences about a speaker's intentions for even a simple question can become complicated. For any particular listener, such inferences rest on an assessment of what the parties involved think about each other—what knowledge, beliefs, or attitudes they believe are shared. This has been called their mutual knowledge (Lewis, 1969; Schiffer, 1972) or *common ground* (Clark and Marshall, 1981). For a belief or piece of information to be a part of two people's common ground, both parties must hold a vast (in principle, infinite) set of reciprocal beliefs: Each person must believe that the other person holds the belief; each must believe that the other believes they hold the belief; each must believe that the other believes that they believe the other holds the belief, etc. (see Clark and Marshall, 1981; Schiffer, 1972).

So, as an addressee you infer your aunt's intentions against the common ground you assume you share with her, along with your assumptions about the common ground she shares with the side participants (the other relatives present). The side participants infer your aunt's intentions against what they assume your common ground with her is, along with their knowledge about their own common ground with her. Bystanders make similar inferences, but with less certainty that the aunt's utterance has been designed with them in mind; eavesdroppers know that the aunt's utterance has not been designed with them in mind. Side participants, bystanders, and eavesdroppers have all been shown to understand the references in utterances less accurately than addressees

(Schober and Clark, 1989; Wilkes-Gibbs and Clark, 1992), and this results in part from the fact that a speaker's utterances are designed most particularly for the addressee.

People can assume common ground with each other on two main bases—cultural and personal. At the cultural level, people can assume that they share relevant kinds of mutual knowledge with other members of the many communities they belong to—English speakers, women, U.S. citizens, parents, New Yorkers, theater aficionados, woodworkers, musicians, etc. At the personal level, people can assume that the physical environment they are currently in is mutually known to both parties. They can also assume that experiences they have shared with the other person are mutually known, including what has been said in current and previous conversations.

Both cultural and personal common ground come into play as you interpret your aunt's question. You interpret your aunt's question knowing that her social role is as a visiting relative, and not as your boss or the IRS; on the basis of your knowledge of social roles, you can assume that her purposes in asking the question don't include using your answer for employing you or levying taxes. If, through earlier personal experience with your aunt, you know that she finds office socializing a burden, you may assume she thinks of "work" in a broadly inclusive way, and you may thus include your lunches with the boss as work hours. If every other question your aunt has pursued thus far during the visit has been about you and your spouse together, you may rely on this personal common ground to infer that her current question probably refers to both of you, and not just yourself.

### 6.1.2 Grounding

Personal common ground accumulates during a conversation through an interactive process that has been called *grounding* (Clark and Brennan, 1991; Clark and Schaefer, 1987, 1989; Clark and Wilkes-Gibbs, 1986). The idea is simple: an utterance can only be said to have entered common ground if both conversational participants ratify that it has been understood. If your aunt tells you that she just returned from a trip to Singapore, she can assume that this is now mutually known—that this utterance is grounded—only if you give her evidence that you understood what she told you. Her utterance alone does not guarantee grounding; you might not have heard her at all, you might not have been paying close enough attention to understand what she said, or you might have misheard her. Only once you provide her with evidence of understanding (or at least no evidence of misunderstanding) can she presume that you have understood her. And only once your aunt presents you with evidence that she has accepted *your* evidence of understanding can you both truly believe that the utterance has entered your common ground.

You can provide your aunt with evidence of understanding in various ways (Clark and Schaefer, 1989), some more explicit than others. You can explicitly acknowledge comprehension by nodding, saying "uh-huh" or "okay," or saying

“I understand.” You can repeat your aunt’s utterance verbatim: “So you went to Singapore!” You can demonstrate that you have understood by continuing the conversation with another utterance that is relevant. Or you can show her implicitly that you understood by continuing to pay attention to her, on the assumption that if you hadn’t understood her you would have asked for clarification.

You can also show your aunt that you *haven’t* understood her utterance by providing explicit or implicit evidence. If you know she said something but you didn’t hear it, you can say “Huh?” or “What was that?” or “I didn’t hear you.” If you only heard part of what she said you can say “You went where?” or “What was that about Singapore?” You can gaze at her uncomprehendingly, implicitly requesting her to repeat or reframe her utterance.

The point is that only once the two of you have agreed that her initial utterance was understood can the utterance be considered grounded. This can take several conversational turns or several clarification sequences. You might provide your aunt with evidence that you have interpreted her utterance in a way that she finds unacceptable (“Oh, your visa expired and you were thrown out?”), and she can work with you to make sure that you understand her statement as she intended it (“No, I didn’t feel like staying any longer”).

It isn’t only statements like “I just got back from Singapore” that must be grounded, but also questions like “So, how many hours a week do you work?” The question is only grounded once both of you agree that it has been understood as intended. If you answer your aunt’s question for both yourself and your spouse (“Probably about 100 hours”) when she only intended to be asking about you, she can persist (“No, I mean just you”) until you finally do interpret her question her way.

Of course, grounding does not guarantee absolutely accurate understanding; it only guarantees that people understand each other as well as they want to for current purposes. In some situations—say, casual cocktail party conversations—people may not care if they understand precise references and exact underlying intentions. In other situations—say, a telephone conversation with someone giving you technical help on using your computer—you may want to be more certain that you have understood. In such cases you will probably use explicit grounding procedures to their fullest extent.

## 6.2 MAKING SENSE OF SURVEY QUESTIONS

Now suppose that you are asked exactly the same question—“How many hours a week do you work?”—in a telephone survey interview administered by a government agency like the Census Bureau or the Bureau of Labor Statistics. Do you make sense of the question in the same way as you do when your aunt asks the question informally? As a respondent in an official survey, you probably know intuitively that you are not engaged in the same kind of conversation as the one with your visiting relatives, even though you may never have thought

about this explicitly. There are clearly different rules of the game at work in the survey interview.

In particular, conversation in standardized surveys is restricted in various ways (see, e.g., Clark and Schober, 1992; Schaeffer, 1991; Suchman and Jordan, 1990, 1992), for all sorts of legitimate historical reasons like concerns about interviewer bias and generalizability (see Beatty, 1995, for a discussion of what has led to current prescriptions for interviewer behavior). Because of these restrictions, audience design and grounding operate differently in standardized surveys than they do in spontaneous conversation.

### 6.2.1 Audience Design

In standardized surveys, interviewers read questions that other people (the survey designers) have scripted at another time and place. The questions are supposed to be read verbatim, with absolutely no deviation. Rather than designing questions on the fly for particular addressees, as your aunt does in her informal conversation with you, survey interviewers are *intermediaries* for the survey designers (see Clark and Schober, 1992). Unlike your aunt, who can freely change the course of her questioning, interviewers are required to follow the script under all circumstances. They must do this even if following the script seems absurd, as when a respondent's elaboration on an answer to one question would ordinarily make the entire set of follow-up questions irrelevant.

Questions in survey interviews aren't designed with particular addressees in mind. Rather, they are scripted (and pretested) to be appropriate for a generic, nonspecific member of the culture. In this respect, audience design in survey interviews resembles the kind of community-based audience design that authors or journalists rely on when they write for the general public, or the kind of audience design that a lecturer or broadcaster uses for speaking to a large audience.

So, if you are savvy about how survey interviews work, you can't presume that the question "How many hours a week do you work?" was designed with exactly your circumstances in mind. It was designed with the average respondent in mind, and to answer the question appropriately you need to imagine how the question could have been intended to be interpretable for the average respondent. If your work situation doesn't match what you guess is a usual situation—you have wildly varying hours, or you perform various activities that may or may not be classifiable as work—you need to figure out how your circumstances map onto the average situation.

You also attribute different motives to the survey designers who created the question than to your aunt, and you have different notions about what your answer will be used for. If you know that the survey you are taking part in is an official government survey that measures unemployment rates, you might assume (correctly or incorrectly) that "work" and "hours" are intended to have particularly stringent interpretations. If you have a fearful streak, you may imagine that your responses will be checked against IRS records for accuracy, and that your answer could get you in trouble for underreported income. If the sur-

vey is a public opinion poll commissioned by a labor advocacy group, or a market research survey, you may assume that the purposes in asking the question are quite different, and you may interpret words in the questions quite differently.

### 6.2.2 Grounding

The grounding procedures available to you and the survey interviewer are severely restricted. As you will notice if you try to ask for clarification (“What exactly do you mean by work?”), standardized interviewers are trained to take a neutral stance in order to avoid suggesting answers or biasing responses in any way. This translates into some very specific procedures for interviewers to adopt (Fowler and Mangione, 1990).

Consider Fowler and Mangione’s prescriptions for what an interviewer should do if you ask for clarification on the work question. For questions like this that demand a numerical response, the interviewer must probe to get you to present one number, but the probe must not bias your response in any way. The permitted methods for doing this include repeating the question (“The question asks: How many hours a week do you work?”), describing what kind of response is needed (“I need a number of hours”) or explicitly saying that your interpretation is required (“What would be *your* best estimate?” or “We need *your* interpretation” or “Whatever it means to you”).

Note how different this is from what your aunt would do. If you ask your aunt “What exactly do you mean by work?” she will probably feel some obligation to tell you what she means, or at least to find out why you want to know and to handle your concern—to ground your understanding. Officially, the survey interviewer is supposed to *avoid* grounding your understanding of the question.

Fowler and Mangione’s other prescriptions also differ from what your aunt is likely to do. For questions that provide response alternatives (“How would you rate your schools—very good, good, fair, or poor?”), if the respondent doesn’t answer with exactly one of these alternatives (as in “The schools around here are not very good”), the interviewer must repeat the question and/or repeat *all* the response alternatives (“Would that be very good, good, fair, or poor?”). Any deviation from this procedure, like only re-presenting some of the response alternatives (“Well, would you say fair or poor?”), would be considered a directive probe (Fowler and Mangione, 1990, pp. 39–40).

Your aunt probably wouldn’t ask your opinion about schools by providing a list of response alternatives. But if she did ask you a question with response alternatives, like “How do you like your oatmeal, thick or runny?” she would be unlikely to re-present all the response alternatives if your answer did not match the ones she had proposed. If you answered, “Oh, fairly thick,” your aunt probably would not say “Would you say thick or runny?” Instead, she would tailor any further inquiries to what you had just said: “Just how thick do you mean?” or “Do you mean really thick or just medium thick?” In other words, unlike an interviewer required to follow standardized procedures, she

would adjust her follow-up question to show that she had understood your answer.

If your aunt were to ask you your opinion about local schools, rather than presenting you with response alternatives, she would probably ask the question in an open-ended way: "How would you rate your schools?" Here is what standardized survey interviewers are licensed to do when responses to open-ended questions are inadequate (Fowler and Mangione, 1990, p. 42). They can:

1. Repeat the question.
2. Probe saying, "How do you mean (that)?"
3. Probe saying, "Could you tell me more about (that)?"
4. Ask, "Is there anything else?"

Your aunt might use some of these interviewer behaviors as you and she ground your understanding of her question, although some of them would sound odd or stilted or interview-like in an informal interaction. But it is highly unlikely that she would restrict herself to *only* these behaviors or that she would try to adopt the neutral stance that they attempt to embody. Rather, she would try to ground understanding by adapting whatever she said to what you said.

In sum, interviewers deviate from ordinary grounding techniques by using some licensed probing techniques. This is especially true for interviewer behaviors that adopt a "neutral" stance, like the "whatever it means to you" probes. On the other hand, interviewers sometimes use ordinary grounding techniques, as when they accept responses by saying "okay" or "uh-huh," when they repeat a question that the respondent didn't hear, and when they give explicit or implicit evidence that they have not understood an answer (see Schober and Conrad, 1997).

Note that grounding in standardized interviews is *asymmetrical*. Respondents can't expect a substantive answer if they ask interviewers what a question means; interviewers aren't supposed to help ground respondents' understanding. But interviewers are licensed to ask respondents what they mean by an answer, and the rules of the game allow—in fact, require—respondents to tell them.

### 6.3 IMPLICATIONS FOR SURVEY RESEARCH

At some level, respondents usually understand that a survey interview is not like an ordinary conversation. Ideally, they would understand the implications of the fact that audience design and grounding only partially operate as they ordinarily do. But I don't believe they always do. I propose that the similarities between grounding techniques in ordinary conversation and survey interviews may sometimes fool respondents into acting as if they were real addressees, rather than participants in an unusual conversation with an intermediary (the



interviewer) for an absent third party (the survey designer or designers). Of course, respondents in survey interviews *aren't* like real addressees, although it is an open question whether they should be considered more like side participants, bystanders, or eavesdroppers. Following Gerrig's (1994) reasoning on the status of readers of fiction, survey respondents are like side participants because questions are written for an audience toward whom the writer has at least *some* intentions. But one could argue that survey respondents are like bystanders and eavesdroppers because, unlike side participants, they cannot interact with the speaker to resolve misunderstandings.

In other words, there are both superficial and substantive ways in which survey interactions appear to be like ordinary conversations (see also Schaeffer, 1991). Superficially, survey interviews and spontaneous conversations both involve two parties exchanging information in a series of turns. More substantively, some of the techniques interviewers use to get answers they can use—saying “uh-huh,” repeating a question the respondent hasn't heard, etc.—overlap with ordinary grounding techniques. Of course, interviewers aren't really supposed to ground meaning; they are supposed to probe neutrally in order to obtain codable, usable answers (see Houtkoop-Steenstra, 1996), ideally without providing any interpretations for the respondent.

This overlap creates an ambiguous situation for survey respondents. Are they real addressees, conversing with a partner who is presenting them with evidence of understanding, or not? When the interviewer says “uh-huh” or “okay,” going on to the next question, does this mean the interviewer has really understood the respondent—with all the concomitant assumptions to be made about what has been grounded—or not?

Such ambiguity, I propose, is at the heart of interviewer-respondent interaction in standardized surveys. It is part of what leads both interviewers and respondents to deviate from official scripts (see Houtkoop-Steenstra, 1996; Schober and Conrad, in press; Suchman and Jordan, 1990, 1992). And it is part of what creates some of the mysterious “response effects” in survey interviews, where question wording, response alternatives, and question ordering can drastically affect the respondents' answers (see also Clark and Schober, 1992; Schwarz, 1994, 1996; Strack and Schwarz, 1992).

### 6.3.1 Response Effects

As a growing body of evidence shows (see, e.g., Bless, Strack, and Schwarz, 1993; Schwarz, 1996, 1998; Strack and Schwarz, 1992; Sudman, Bradburn, and Schwarz, 1996), survey respondents don't necessarily switch off their ordinary conversational reasoning. Although survey respondents shouldn't depend on their ordinary assumptions about audience design and their ordinary procedures for grounding as they make sense of survey questions, it seems that sometimes they do.

An exhaustive overview is beyond the scope of this chapter, but a few examples can help make the point. In each case, the oddness of audience design and

grounding in the survey situation helps explain the seemingly anomalous survey results.

**Example 1: Response Alternatives** Several studies have shown that the response alternatives presented as part of a question influence respondents' answers. In one study (Schwarz, Hippler, Deutsch, and Strack, 1985), German respondents were asked to report how much television they watched per day. One set of respondents were presented with response alternatives in a low range, from "up to a half hour" to "more than two and a half hours." The other set of respondents were presented with the same number of response alternatives in a high range, from "up to two and a half hours" to "more than four and a half hours." When asked the question with the low range of response alternatives, 16.2 percent of the respondents reported watching more than two and a half hours of television per week. In contrast, 37.5 percent of the respondents given the high range of response alternatives reported watching that much television.

How does audience design help explain this? Respondents assume—perhaps mistakenly—that the question wording (in this case, the response scale intervals) is informative about what the survey researchers intended. The survey researchers who designed the questions presumably knew something about national TV-watching habits, and so the response scales probably represent what is normal within the population. Indeed, later questioning showed that respondents with the high range estimated that Germans watch more television (3.2 hours per week) than respondents given the low range (2.7 hours per week). Respondents mistakenly took the distribution of the response scales as based on normal TV-watching behavior, because they presumed the survey designers were knowledgeable.

Note that this audience design operates through cultural and not personal common ground. That is, these respondents aren't assuming that the response alternatives were designed for them personally based on what the researchers know about them, or based on what they have told the interviewer thus far. Instead, respondents are judging what the interviewer must have meant through the wording of the question itself.

In ordinary conversation, addressees regularly make sense of questions through response alternatives. If your aunt asks "Do you prefer red or white wine?" you will probably assume that these are your only choices (possibly this is all she has in the house), and that rosé is not among the alternatives. This is because the response alternatives she provided have restricted the *domain of inquiry* (see Clark and Schober, 1992) and set the presuppositions for asking the question.

So it isn't surprising that response scales affect more than just frequency estimates in surveys. They also affect respondents' reports of their subjective experience (e.g., Schwarz, Strack, Müller, and Chassein, 1988). Other differences in response alternatives also affect people's answers, as when a set of

response alternatives includes or excludes “don’t know” as an option (see Clark and Schober, 1992, for a review).

It is also not surprising that effects of response alternatives can be reduced or eliminated entirely. This can be done if researchers explicitly block the basis for respondents’ assuming that the response alternatives are informative. For example, as Schwarz (1996) reports, when respondents are told that they are participating in the study of German TV-watching so that the researchers can determine what the right response alternatives are, the effects of the response alternatives disappear.

**Example 2: Question Ordering** Even though survey researchers often intend each question in a survey to be answered independently of all the others, respondents sometimes provide different answers depending on the order the questions appear in (see, for example, Hippler and Schwarz, 1987; Hyman and Sheatsley, 1950; Rugg and Cantril, 1944; Schuman and Presser, 1981; Strack, Martin, and Schwarz, 1988). Here is one particularly clear example (Schwarz, Strack, and Mai, 1991).

In the study, respondents were asked several questions (in German) to be answered on an 11-point rating scale, from very dissatisfied to very satisfied. One question (A) was about general life satisfaction: “How satisfied are you currently with your life as a whole?” Another question (B) was about relationship satisfaction: “Please think about your relationship to your partner (spouse or date). How satisfied are you currently with your relationship?” When question A preceded question B, the answers were significantly less highly correlated ( $r = 0.32$ ) than when question A followed question B ( $r = 0.67$ ).

How might audience design and grounding help account for this? In ordinary conversation, addressees presume that questions are designed based not only on cultural common ground but also on personal common ground. When your aunt asks you a question after you have just answered an earlier question, you assume that the content of your answer was grounded. The mere fact that she asked you another question on a different topic provides evidence that she understood your answer well enough for her current purposes, and that your answer entered your personal common ground with her. So her next question, you presume, is being asked *in light of* your answer to the first question, and you will interpret the question accordingly.

So if your aunt asks you how satisfied you are with your current relationship and then asks you about your happiness with life in general, you will assume that her intention is something like this: “Given what you just told me about your current relationship, how happy are you with life in general?” In other words, you will assume that the second question is asking for elaboration on your first response, and that your aunt assumes the two areas of your life are connected. On the other hand, if your aunt asks you about your general happiness and then asks about your satisfaction in your current relationship, you will assume her intention is something like this: “You just told me how happy you

are in general. Now I'm asking about a more particular aspect of your life." In this case, you are less likely to assume that the second question builds on the first, because your evaluation of general life happiness may include many other aspects of life.

This interpretation is supported by some additional results. Schwarz, Strack, and Mai reworded the questions to explicitly suggest the inclusive interpretation, by starting off with "Including the life-domain that you already told us about, how satisfied are you currently with your life as a whole?" This led the two answers to be highly correlated again ( $r = 0.61$ ), just about as highly as they had been correlated when they had simply been presented in that order. Schwarz et al. also altered the wording to make explicit that the two questions should not be interpreted as related ("We would first like to ask you to report on two aspects of your life, which may be relevant to people's overall well-being: (a) relationship satisfaction; (b) satisfaction with life as a whole."). This led the two answers to be far less correlated ( $r = 0.18$ ), which is as low as the two answers were correlated when the instructions specifically stated that the second question was about a different area entirely ( $r = 0.20$ ).

So unless they are explicitly informed that they shouldn't, respondents interpret subsequent questions in light of previous ones, just as they do in ordinary conversation. Now, there are also cognitive processes like priming involved in question order effects (see, e.g., Sudman et al., 1996, for a discussion), and these may not depend on interactive processes like audience design or grounding. But the point is that interactive processes can be involved: respondents don't shut off their ordinary assumptions about what has been grounded, even though they might be aware that interviewers are following a script written at an earlier time by survey designers who aren't there to ground understanding.

**Example 3: Differing Interpretations of Questions** Different respondents can interpret exactly the same questions in radically different ways. We know this from comparisons of survey responses to official records, like police reports, bank statements, and hospital records (see Wentland, 1993, for a review). We also know it from subsequent questioning of respondents, when they are asked, after they have participated in a survey, why exactly they answered the questions as they did.

To take an extreme (and controversial) example of a finding using this second technique, Belson (1981, 1986) reported that a notable percentage of respondents interpreted even ordinary words and phrases in survey questions—words like "weekend," "you," "children," and "generally"—differently than the question designers had intended. For example, when respondents were asked "For how many hours do you usually watch television on a weekday? This includes evening viewing," 15 percent of the respondents interpreted "you" as including other people too, 33 percent included times the television was on but they hadn't been paying attention to it, and 61 percent included days other than the five weekdays intended or excluded some weekdays for other reasons (see Belson, 1981, pp. 127–137).

Why might this be? Part of the problem is that alternate interpretations are always available for words in ordinary conversation. Even though addressees presume that speakers have used audience design in selecting their words, this doesn't guarantee perfect understanding. When your aunt asks about your work hours, you have to make a judgment about who exactly "you" includes. Since you have the ability to ground understanding, you can ask your aunt exactly what she means by "you." You also know that if your interpretation of "you" turns out to have been different from your aunt's intention, the mistake can be corrected, if you and your aunt so desire.

In standardized survey interviews, respondents are not licensed to ground understanding of words and phrases in the questions. Respondents must answer questions with a *presumption of interpretability* (Clark and Schober, 1992)—they have to assume that their best guess as to an interpretation must be the one the survey designers intended. And their best guess can easily be wrong.

Evidence for this explanation—that the inability to ground understanding is one major factor leading to inaccurate interpretations—can be found in some studies that Fred Conrad and I have carried out. In two studies, one laboratory study (Schober and Conrad, 1997) and one field study using a national telephone sample (Schober and Conrad, 1998), we have compared respondents' answers to the same questions administered by interviewers who either (a) follow strict standardized interviewing procedures or (b) follow more conversationally flexible procedures. Respondents answered fact-based questions from major government surveys about housing (e.g., "How many bedrooms are there in your house or apartment?"), jobs (e.g., "Last week, did you have more than one job, including part-time, evening, and weekend work?"), and purchases (e.g., "In the past year, did you purchase or have expenses for household furniture?").

In both studies, we trained one group of experienced telephone interviewers to adhere to Fowler and Mangione's (1990) prescriptions for standardization. They were to take a neutral stance, never providing any definitions or interpretations of questions for respondents. We trained another group to help the respondents understand the questions in the way the survey designers had intended—that is, according to the official definitions for key words and phrases in the questions that the survey organizations had provided. Specifically, this second group of interviewers were allowed to rephrase questions, answer requests for clarification, and provide respondents with additional information or clarification even if they had not directly asked for it.

In the laboratory study, respondents answered questions on the basis of fictional scenarios we had devised—floor plans of houses, descriptions of work situations, or purchase receipts. Thus, we knew what the facts were and could code response accuracy—that is, whether responses matched official definitions. In the field study, we couldn't measure response accuracy directly, because we didn't know people's living and work situations. So we relied on indirect measures of response accuracy: how often their responses in a second interview (flexible or standardized) differed from their responses to the same question in

an initial standardized interview, and how often the purchases they listed when they answered “yes” to a purchase question conformed to the official definitions.

In both studies, flexible interviewing techniques led respondents to produce answers that were more consistent with the official definitions. This was especially the case when the mapping between respondents’ circumstances and the official definition wasn’t obvious to the respondents—for example, when respondents weren’t sure whether to classify a television purchase as a household furniture purchase or not (the official government definition excludes televisions as furniture).

What these results show is that the match between survey designers’ intentions and respondents’ interpretations of questions can be quite poor. Standardized interviewing techniques, which do not allow respondents to ground their understanding of questions, can lead to poor response accuracy, exactly because this match is poor. More flexible interviews, which allow interviewers and respondents to ground the respondent’s understanding, can improve the match, thus leading to more accurate responses.

But the improvement in response accuracy resulting from licensed grounding comes at a real cost: increased interview duration. These results should not be seen as promoting large-scale adoption of flexible interviewing techniques; far too much remains to be seen about the potential costs and benefits of the technique. The point here is that if we conceive of survey interviews as interactional phenomena, in which some unusual versions of audience design and grounding are taking place, we can study new theoretically-based ways to improve response accuracy.

## 6.4 CONCLUSIONS

Respondents make sense of survey questions in ways that are related to the ways they make sense of questions as addressees in ordinary conversation. They make inferences about the intentions and agendas underlying the questions—how they presume the question designers implemented *audience design*. This amounts, in part, to judgments about the common ground, both cultural and personal, that the question designers may have taken into account as they designed the survey.

Beyond these individual judgments that respondents make, they also rely on interactive *grounding* procedures to make sense of questions, although the procedures available to them in surveys only overlap partially with the procedures available in ordinary conversation. Respondents sometimes seem to treat surveys as if grounding is proceeding in the ordinary way; this can help account for certain question order effects in surveys. The inability of survey respondents to fully ground their understanding of questions can lead to wide variability in how respondents interpret particular words and phrases in survey questions.

I should note that standardized practice varies from survey to survey, and that not everyone follows Fowler and Mangione’s (1990) prescriptions to the

letter. It remains to be seen what the precise effects of particular deviations are. But in examining such effects, an interactional analysis of the type proposed here is a good place to start.

## ACKNOWLEDGMENTS

This work was supported in part by NSF grant IRI-94-02167. Many thanks to Herbert Clark, Fred Conrad, and Norbert Schwarz for comments on earlier versions of this paper. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

## REFERENCES

- Beatty, P. (1995). Understanding the standardized/non-standardized interviewing controversy. *Journal of Official Statistics*, 11, 147-160.
- Belson, W. A. (1981). *The Design and Understanding of Survey Questions*. Aldershot, England: Gower.
- Belson, W. A. (1986). *Validity in Survey Research*. Brookfield, VT: Gower.
- Bless, H., Strack, F., and Schwarz, N. (1993). The informative functions of research procedures: Bias and the logic of conversation. *European Journal of Social Psychology*, 23, 149-165.
- Clark, H. H. (1992). *Arenas of Language Use*. Chicago: University of Chicago Press.
- Clark, H. H. (1996). *Using Language*. Cambridge, UK: Cambridge University Press.
- Clark, H. H., and Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, and S. D. Teasley (Eds.), *Perspectives on Socially Shared Cognition*, pp. 127-149. Washington, DC: APA.
- Clark, H. H., and Carlson, T. B. (1982). Hearers and speech acts. *Language*, 58, 332-373.
- Clark, H. H., and Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. Webber, and I. Sag (Eds.), *Elements of Discourse Understanding*, pp. 10-63. Cambridge, UK: Cambridge University Press.
- Clark, H. H., and Schaefer, E. F. (1987). Collaborating on contributions to conversations. *Language and Cognitive Processes*, 2, 19-41.
- Clark, H. H., and Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, 13, 259-294.
- Clark, H. H., and Schober, M. F. (1992). Asking questions and influencing answers. In J. M. Tanur (Ed.), *Questions About Questions: Inquiries into the Cognitive Bases of Surveys*, pp. 15-48. New York: Russell Sage Foundation.
- Clark, H. H., and Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1-39.

- Fowler, F. J., and Mangione, T. W. (1990). *Standardized Survey Interviewing: Minimizing Interviewer-Related Error*. Newbury Park, CA: Sage.
- Gerrig, R. (1994). *Experiencing Narrative Worlds*. New Haven: Yale University Press.
- Goffman, E. (1981). *Forms of Talk*. Philadelphia: University of Pennsylvania Press.
- Graesser, A. C., and Franklin, S. P. (1990). QUEST: A cognitive model of question-answering. *Discourse Processes*, 13, 279–303.
- Grice, H. P. (1957). Meaning. *Philosophical Review*, 66, 377–388.
- Grice, H. P. (1975). Logic and conversation. In P. Cole and J. L. Morgan (Eds.), *Syntax and Semantics: 3. Speech Acts*, pp. 41–58. New York: Academic Press.
- Hippler, H. J., and Schwarz, N. (1987). Response effects in surveys. In H. J. Hippler, N. Schwarz, and S. Sudman (Eds.), *Social Information Processing and Survey Methodology*, pp. 102–122. New York: Springer-Verlag.
- Houtkoop-Steenstra, H. (1996). Probing behaviour of interviewers in the standardised semi-open research interview. *Quality and Quantity*, 30, 205–230.
- Hyman, H. H., and Sheatsley, P. B. (1950). The current status of American public opinion. In J. C. Payne (Ed.), *The Teaching of Contemporary Affairs: Twenty-First Yearbook of the National Council of Social Studies*, pp. 11–34. New York: National Education Association.
- Lewis, D. (1969). *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.
- Rugg, D., and Cantril, H. (1944). The wording of questions. In H. Cantril (Ed.), *Gauging Public Opinion*, pp. 23–50. Princeton, NJ: Princeton University Press.
- Schaeffer, N. C. (1991). Conversation with a purpose—or conversation? Interaction in the standardized interview. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, and S. Sudman (Eds.), *Measurement Errors in Surveys*, pp. 367–391. New York: Wiley.
- Schiffer, S. (1972). *Meaning*. Oxford, UK: Clarendon Press.
- Schober, M. F. (1998). Different kinds of conversational perspective-taking. In S. R. Fussell and R. J. Kreuz (Eds.), *Social and Cognitive Approaches to Interpersonal Communication*, pp. 145–174. Mahwah, NJ: Erlbaum.
- Schober, M. F., and Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, 21, 211–232.
- Schober, M. F., and Conrad, F. G. (in press). A collaborative view of standardized surveys. To appear in J. van der Zouwen, N. C. Schaeffer, D. Maynard, and H. Houtkoop (Eds.), *Interviewer-Respondent Interaction in the Standardized Survey Interview*.
- Schober, M. F., and Conrad, F. G. (1997). Does conversational interviewing reduce survey measurement error? *Public Opinion Quarterly*, 60, 576–602.
- Schober, M. F., and Conrad, F. G. (1998). Does conversational interviewing improve survey data quality beyond the laboratory? *1997 Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 910–915.
- Schuman, H., and Presser, S. (1981). *Questions and Answers in Attitude Surveys: Experiments in Question Form, Wording and Context*. New York: Academic Press.



- Schwarz, N. (1994). Judgment in a social context: Biases, shortcomings, and the logic of conversation. In M. Zanna (Ed.), *Advances in Experimental Social Psychology*, Vol. 26, pp. 123–162. San Diego, CA: Academic Press.
- Schwarz, N. (1996). *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation*. Mahwah, NJ: Erlbaum.
- Schwarz, N. (1998). Communication in standardized research situations: A Gricean perspective. In S. R. Fussell and R. J. Kreuz (Eds.), *Social and Cognitive Approaches to Interpersonal Communication*, pp. 39–66. Mahwah, NJ: Erlbaum.
- Schwarz, N., Hippler, H. J., Deutsch, B., and Strack, F. (1985). Response scales: Effects of category range on reported behavior and comparative judgments. *Public Opinion Quarterly*, 49, 388–395.
- Schwarz, N., Strack, F., and Mai, H. P. (1991). Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis. *Public Opinion Quarterly*, 55, 3–23.
- Schwarz, N., Strack, F., Müller, G., and Chassein, B. (1988). The range of response alternatives may determine the meaning of the question: Further evidence on informative functions of response alternatives. *Social Cognition*, 6(2), 107–117.
- Strack, F., Martin, L. L., and Schwarz, N. (1988). Priming and communication: The social determinants of information use in judgments of life-satisfaction. *European Journal of Social Psychology*, 18, 429–442.
- Strack, F., and Schwarz, N. (1992). Communicative influences in standardized question situations: The case of implicit collaboration. In G. R. Semin and K. Fiedler (Eds.), *Language, Interaction and Social Cognition*, pp. 173–193. London: Sage.
- Suchman, L., and Jordan, B. (1990). Interactional troubles in face-to-face survey interviews. *Journal of the American Statistical Association*, 85(409), 232–241.
- Suchman, L., and Jordan, B. (1992). Validity and the collaborative construction of meaning in face-to-face surveys. In J. M. Tanur (Ed.), *Questions About Questions: Inquiries into the Cognitive Bases of Surveys*, pp. 241–267. New York: Russell Sage Foundation.
- Sudman, S., Bradburn, N., and Schwarz, N. (1996). *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass.
- Wentland, E. J. (1993). *Survey Responses: An Evaluation of Their Validity*. San Diego, CA: Academic Press, Inc.
- Wilkes-Gibbs, D., and Clark, H. H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language*, 31, 183–194.