CHAPTER ONE

# An Introduction and a Point of View

Survey research rests on the age-old practice of finding things out by asking people questions. In this respect, it has much in common with a diverse set of activities ranging from police interrogations and courtroom proceedings to medical interviews and quiz shows. At the heart of each situation, one person asks another person questions for the purpose of obtaining information (Schuman & Presser, 1981).

In surveys, this pair is the *interviewer* and the *respondent*. Interviewers can put questions to respondents face-to-face, over the telephone, or through a computer. However, the interviewer's questions and the respondent's answers are always the central ingredients. The questions can ask about the personal activities or circumstances of the respondent (often called *behavioral* or *factual questions*) or they can seek the respondent's opinion about an issue (*attitude questions*). The examples in (1) are factual questions, and those in (2) are attitude questions from national surveys:

(1) a. Was anything stolen from you while you were away from home, for instance at work, in a theater or restaurant, or while traveling? [NCS][1]

b. Since the 1st of (month, 3 months ago), have you (or any members of your C[onsumer] U[nit]) received any bills for telephone services? Do not include bills used entirely for business purposes. [CE]

[1] In citing examples from surveys, we follow the typographic conventions of the surveys themselves. We indicate the source of the survey in brackets after the example, using the following abbreviations for the five surveys from which we draw most of our examples: CE for Consumer Expenditure Survey, CPS for Current Population Survey, GSS for General Social Survey, HIS for the U.S. Health Interview Survey, and NCS for National Crime Survey.



Copyrighted material

(2) a. In general, do you favor or oppose the busing of (Negro/Black) and white school children from one school district to another? [GSS]
b. Everything considered, would you say, in general, you approve or disapprove of wiretapping? [GSS]

The factual/attitude terminology fits some survey questions more comfortably than others, but these examples give a rough idea of the types of questions that will concern us here.

One might suppose that survey researchers would long ago have developed detailed models of the mental steps people go through in answering survey questions, models that spell out the implications these steps have for survey accuracy. After all, the accuracy of surveys depends almost completely on the accuracy of people's answers. Of course, with attitude questions, it is often difficult to decide what an accurate answer is. Still, answers to such questions are prone to a variety of well-documented *response effects* – differences in survey outcomes that reflect seemingly irrelevant procedural details such as the order in which the answer categories are presented. These response effects may be due to problems in understanding the question, remembering relevant information, producing an appropriate answer, or other mental processes. Despite their importance for understanding and evaluating surveys, the study of the components of the survey response process is in its infancy, having begun in earnest only in the 1980s.

By contrast, rigorous mathematical formulations of *sampling errors* have been available since the early 1950s, when the major texts on survey sampling appeared (Cochran, 1953; Deming, 1950; Hansen, Hurwitz, & Madow, 1953). Sampling error arises because it is usually impossible to interview everyone in the target population; thus, there is some uncertainty about the relation between estimates that come from the interviewed respondents and the answers that might have come from the population as a whole. The study of survey sampling is devoted to reducing errors attributable to such effects, and it is both precise and scientific. But the study of questionnaire design, whose aim is to reduce response effects, remains an art (as Sheatsley, 1983, and Sudman & Bradburn, 1982, have noted).

A major goal of this book is to provide a framework for understanding response effects in surveys. We will propose a model of the survey response process and trace its implications for survey error. Our hope is to unify what is known about response effects by relating existing find-

Copyrighted material

ings to a cognitive model of the survey response process. Many reviews of the literature on response effects in surveys have noted the fragmentary character of the research on this topic. We believe that a cognitive model of the survey response process can offer an improved map of this varied landscape.

The terms *response effects*, *response process*, and *respondent* misleadingly suggest a behavioristic focus on the way that people respond physically to survey questions. We will continue to use these terms because they are universal in the survey literature, but we have already indicated that we do not share this point of view. Errors on surveys can be due to internal features of language comprehension, memory, and choice, as well as to the way people execute a response, and our theory builds on current research in cognitive psychology and artificial intelligence. We view response effects in surveys as a challenging test case for cognitive science, one that goes beyond the simple tasks that typically find their way into the cognitive laboratory. In studying survey errors, we follow Newell's (1973) advice to consider large-scale domains that draw on many mental abilities and that can therefore shed light on how cognition works as a unified system. We believe (and hope to persuade you) that a cognitive model of how people answer survey questions can offer as many insights to cognitive and social psychologists as to survey methodologists.

## 1.1 Earlier Theories of the Response Process

Before presenting our own view of how people answer survey questions, we consider two important antecedents: psychometric theories of attitude measurement and early proposals within the survey methodology tradition. Comparing these theories to our own approach highlights what's distinctive to the point of view presented in this book.[2]

[2] We concentrate in this section on historical antecedents, but contemporary research in artificial intelligence (AI) and linguistics has also influenced our thinking about the way people answer survey questions, as will be evident in later chapters. For instance, work in AI has examined the process by which computers can be programmed to answer questions framed in everyday language (e.g., Allen, 1995; Graesser, McMahen, & Johnson, 1994; Lehnert, 1978). The AI and linguistics research has concentrated on the difficult problem of identifying the specific information that ordinary questions seek. So far, however, this research has had little direct impact on models of the survey response process. We attempt to correct this oversight in Chapter 2.

Copyrighted material

### 1.1.1 Psychometric Theories

Perhaps the first models of the survey response process were the ones developed by psychometricians – Guttman, Guilford, Likert, and their colleagues. This research produced the first systematic attitude measurement techniques, techniques that quantified the strength of a person's conviction in an opinion (e.g., *How strongly do you favor increasing aid to the homeless?*). Of these early pioneers, Louis Thurstone was perhaps the leader in describing the underlying psychological processes that made attitude measurement possible (see, e.g., Thurstone, 1927).[3] Thurstone's goal was to develop mathematical models that described the outcome of comparisons among several stimuli, including several statements about an issue (e.g., *Aid to the homeless should be increased; Society already does enough for the homeless; The homeless are a nuisance*).

Thurstone's formal models rested on a psychological theory of judgment: the idea that people represent stimuli as points or regions on an internal dimension (e.g., the dimension of strength of agreement with a position). Thus, his psychological models distinguished several component processes, including the identification of the dimension of judgment (or, as Thurstone referred to it, the *psychological continuum*), the judge's reaction to the stimulus (the *discriminal process*), the assignment of scale values, and the comparison of pairs of stimuli (the calculation of their *discriminal difference*). In the case of an attitude question, Thurstonian theory predicts that people compare their own position with what they take to be the position implied by each statement in the scale. The result of the comparison determines which statements they endorse.

For a variety of reasons, these psychometric models have not played a prominent role in discussions of survey error. One major problem is that, even though their emphasis is statistical, the psychometricians' mathematical models are not easily translated into the terms that have come to dominate discussions of survey error; the survey error models are extensions of those used to analyze random sampling error, and they focus on factual questions, for which it is possible (at least in principle) to measure the accuracy of the answer (e.g., Hansen et al., 1953; see Chapter 11 and Groves, 1989, on the differences between the survey and psychometric approaches to measurement error). In addition, from our

[3] Perhaps another antecedent can be found in Frank Ramsey's (1931) work on measurement of utility, although this approach did not join the mainstream in psychology until the 1950s.

Copyrighted material

current vantage point, the earlier psychometric models are hamstrung by the absence of a detailed description of the judgment process. Although Thurstone's model presents one view of how people choose among response options, it is silent on questions of how they identify the psychological continuum on the basis of the survey question, how they recruit relevant attitude information from memory, and many other crucial issues.

### 1.1.2 Cannell's Process Theory

The model proposed by Cannell, Miller, and Oksenberg (1981; see Cannell, Marquis, & Laurent, 1977, for an earlier version) was perhaps the first model of the survey response process to reflect the new cognitive outlook within psychology. Their model distinguished two routes to an answer, one based on relatively careful processing of the question and the other based on superficial features of the interview situation, such as the interviewer's appearance (see Cannell et al., 1981). Careful answers are, according to this model, the product of the five sets of processes in (3):

(3) a. Comprehension of the question;
b. Cognitive processing (i.e., assessments concerning the information sought, retrieval of relevant memories, and integration and response formulation);
c. Evaluation of the accuracy of the response;
d. Evaluation of the response based on goals other than accuracy;
e. Accurate responding.

We illustrate the arrangement of these steps in Figure 1.1. Cannell, Miller, and Oksenberg (1981) view these five sets of processes as more or less sequential, although their model explicitly allowed for the possibility that respondents might cycle back to an earlier stage if they judged their preliminary answer not to be accurate enough. Prior to the fifth stage, when respondents give an accurate (or at least adequate) answer, respondents could switch to the parallel track and alter their answer based on relatively superficial cues available in the interview situation – cues such as the interviewer's appearance or the implied direction of the question. Responses based on such cues were likely to be biased by *acquiescence* (the tendency to agree) or *social desirability* (the need to present oneself in a favorable light).

The model by Cannell and his colleagues has many attractive features,

Copyrighted material

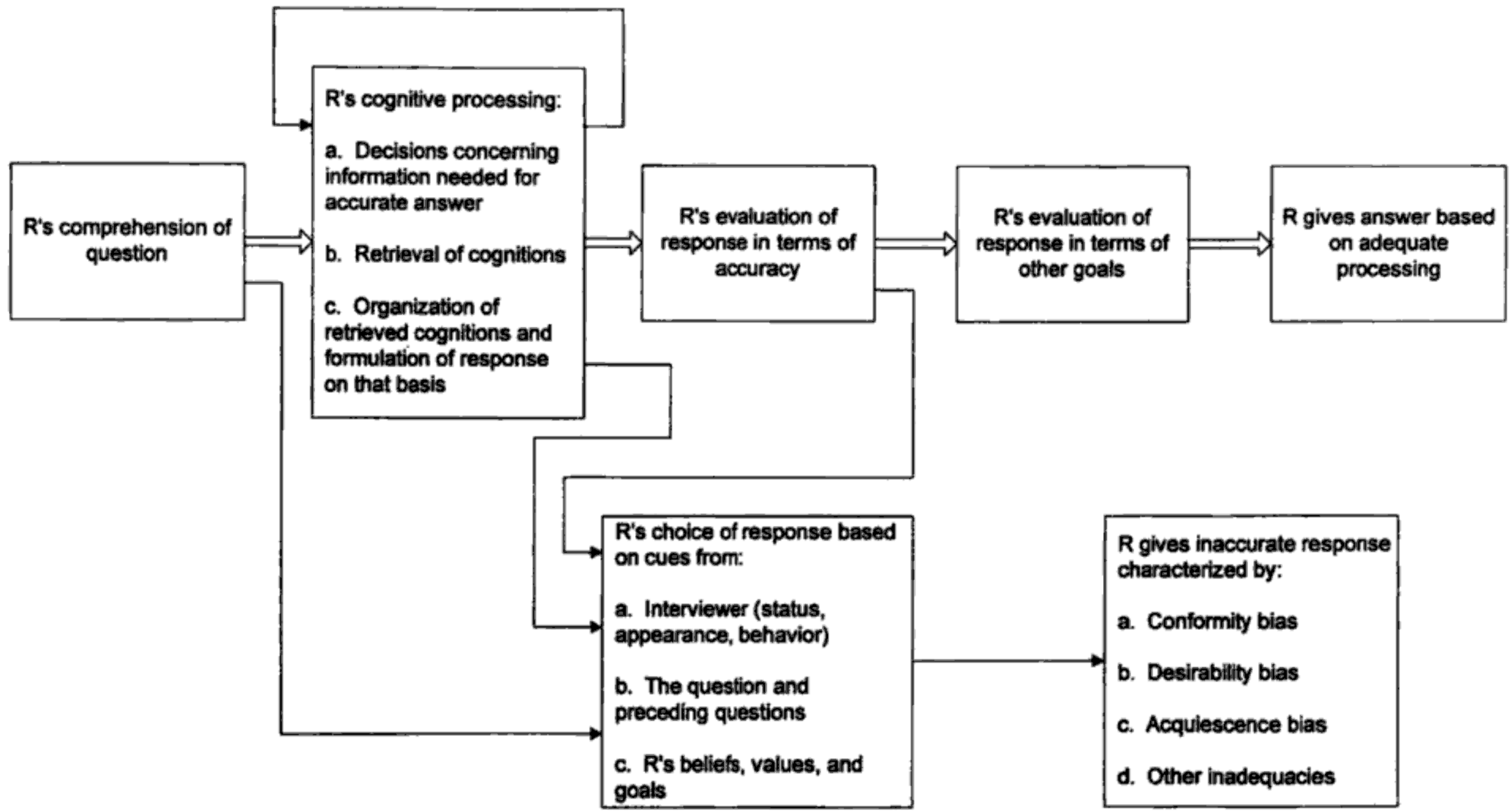


**Figure 1.1.** Cannell, Miller, and Oksenberg's model of the survey response process. The boxes along the bottom of the figure represent the processes needed for a careful answer; those along the top represent the processes leading to an inadequate answer. Adapted from Cannell et al. (1981). Copyright © 1981. Adapted with permission of Jossey-Bass.

Copyrighted material

and it has spawned many related approaches (see Section 1.3). The notion that respondents might take different routes to arrive at an answer is an appealing one, and the specific routes in the Cannell model – one based on systematic processing of the question and the other based on more superficial processing – have a number of parallels in the psychological literature. For example, discussions of attitude change have identified central and peripheral routes to persuasion (Chaiken, 1980; Petty & Cacioppo, 1984). Hastie and Park's (1986) distinction between *memory-based judgments* and *on-line* (i.e., situation-based) *judgments* also bears similarities to the two tracks in Figure 1.1. Another attractive feature of the model is its explicit concern with the respondent's motivation, including such motives as the desire to provide accurate information, to appear agreeable, and to avoid embarrassment.

From our viewpoint, the model suffers from two major drawbacks. The first is that, because the model never assumed a central place in Cannell's work, it was never worked out it much detail. The most complete exposition of the model (in Cannell et al., 1981) runs no more than three pages. Most of the research inspired by the model has focused on improving respondent motivation in a general way rather than on testing predictions regarding the model's specific components. The second drawback is related to the first. It is the model's rather sketchy treatment of the cognitive processes involved in responding to a question, which the model's second stage lumps together. By contrast, the model distinguishes several stages in describing what happens *after* the respondent derives a preliminary answer. The respondent evaluates the initial answer in terms of its accuracy, then in terms of its compatibility with other goals, and finally may modify or discard it based on these earlier assessments. The model seems to assume that respondents could answer questions accurately if only they wanted to and concentrates on whether they decide to answer accurately or not. We favor a different emphasis.

## 1.2 A Proposed Model of the Response Process

We have organized this book around a model that divides the survey response process into four major components – comprehension of the item, retrieval of relevant information, use of that information to make required judgments, and selection and reporting of an answer (Tourangeau, 1984, 1987; Tourangeau & Rasinski, 1988; see also Strack & Martin, 1987). Table 1.1 lists each of these components along with

Copyrighted material

TABLE 1.1 Components of the Response Process

| Component | Specific Processes |
|---|---|
| Comprehension | Attend to questions and instructions<br>Represent logical form of question<br>Identify question focus (information sought)<br>Link key terms to relevant concepts |
| Retrieval | Generate retrieval strategy and cues<br>Retrieve specific, generic memories<br>Fill in missing details |
| Judgment | Assess completeness and relevance of memories<br>Draw inferences based on accessibility<br>Integrate material retrieved<br>Make estimate based on partial retrieval |
| Response | Map judgment onto response category<br>Edit response |

specific mental processes that they might include. In describing these processes, we don't mean to suggest that respondents necessarily perform them all when they answer a survey question. Although some processes may be mandatory, others are clearly optional – a set of cognitive tools that respondents can use in constructing their answer. Exactly which set of processes they carry out will depend on how accurate they want their answer to be, on how quickly they need to produce it, and on many other factors. In this respect, the theory presented in Table 1.1 resembles approaches to decision making that emphasize the array of strategies that people bring to bear on a problem (Payne, Bettman, & Johnson, 1993). These processes are also not exhaustive. We suggest some additions to the list in later chapters when we take up the components in more detail.

Each of these components can give rise to response effects; respondents may, for example, misinterpret the question, forget crucial information, make erroneous inferences based on what they do retrieve, or map their answers onto an inappropriate response category. Both psychological and survey research provide ample evidence of the errors each component produces, and to understand how the response process can go awry, we need to take a closer look at them. In reviewing these components, we also preview the material we will cover in the rest of this book.

Copyrighted material

### 1.2.1 Comprehension

Comprehension encompasses such processes as attending to the question and accompanying instructions, assigning a meaning to the surface form of the question, and inferring the question's point – that is, identifying the information sought (Clark, 1985; Graesser, McMahen, & Johnson, 1994; Lehnert, 1978). Current research on the pragmatics of natural language emphasizes the reasoning people must perform in order to grasp the full implication of a sentence (Grice, 1989; Lewis, 1979; Sperber & Wilson, 1986); research on conversation emphasizes the way conversational partners cooperate to shape their mutual understanding (Clark & Schober, 1992).

As survey researchers have known for some time, many reporting problems arise because respondents misunderstand the questions. Respondents' attention may wander during an interview, and they may miss part of the question; in a self-administered questionnaire, they may not notice essential instructions or, having noticed, they may not bother to read them. The question may be double-barreled, inadvertently asking two or more questions at the same time. The question may include terms that are unfamiliar to the respondent or terms that are understood in different ways by different respondents. The question may be too complicated syntactically or it may contain detailed qualifications that are hard to understand. Familiar terms may nonetheless be vague, and even seemingly clear categories (such as *siblings*) may include borderline cases that can be misclassified (stepbrothers). With attitude questions, a key step in the comprehension process is identifying what issue the question is about. Chapter 2 of this book examines the comprehension of survey questions in more detail; it presents a more detailed theory of this component and reviews what is known about comprehension problems in surveys.

### 1.2.2 Retrieval for Factual Questions

The retrieval component involves recalling relevant information from long-term memory. This component encompasses such processes as adopting a retrieval strategy, generating specific retrieval cues to trigger recall, recollecting individual memories, and filling in partial memories through inference. Several characteristics of the recalled material and of the initiating questions can affect the accuracy and the completeness of this component (Jobe, Tourangeau, & Smith, 1993). These include the

Copyrighted material

distinctiveness of the events, the degree of fit between the terms used in the question and the events' original encoding, the number and quality of the cues that the question provides, the source of the memory (direct experience or secondhand knowledge), and the length of time since the events occurred. Chapter 3 examines these issues.

### 1.2.3 Judgment for Factual Questions

Retrieval often does not yield an explicit answer to survey questions. The question may concern the total number of visits to the doctor in the last six months, the number of hours worked, or the total amount spent on retail purchases during the reference period. If so, respondents must sum the individual events they recalled during the retrieval phase in order to find the total number. The judgment component comprises the processes that respondents use to combine or supplement what they have retrieved. There are at least five major types of judgment processes that may come into play:

(4) a. Judgments regarding the completeness or accuracy of retrieval;
b. Inferences based on the process of retrieval;
c. Inferences that fill in gaps in what is recalled;
d. Integration of the products of retrieval into a single overall judgment;
e. Estimates that adjust for omissions in retrieval.

The first three types of judgment depend on the relation between judgment and retrieval. Type (4a) determines whether further retrieval is warranted and whether specific memories fall within the scope of the question. Some judgments of this type can be seen as an extension of the comprehension component as respondents attempt to implement their understanding of the nature of the events covered by the question. The second type of judgment involves drawing conclusions from features of the retrieval process; for example, when retrieval is difficult or sketchy, respondents may conclude that the events in question happened infrequently or long ago or never took place at all (Brown, Rips, & Shevell, 1985; Gentner & Collins, 1981; Tversky & Kahneman, 1973). The judgments of the third type are attempts to reconstruct what happened by inferring missing details, often based on what typically happens during an event of a given type.

Respondents undertake the remaining types of judgment to transform the retrieved information into an appropriate answer. The fourth process is necessary because people often retrieve fragmentary information that

Copyrighted material

they must combine to produce a single response. This combination may involve simple numerical averaging (Anderson, 1981) or more complex types of estimation. The final process involves situations in which what is remembered forms part of a larger estimation strategy. For example, to answer a question about a lengthy time period, respondents may recall the number of events in a recent, easily remembered portion of the period and then extrapolate to cover the entire period.

The processes in (4) usually operate on retrieved information, but sometimes judgment supplants retrieval entirely. As Reder (1987) has argued, people sometimes answer retrospective questions by considering the general plausibility of a response. Respondents who have never heard the term *health maintenance organization* (*HMO*) can probably infer from that fact alone that they do not belong to one. People can bypass retrieval of specific information in such cases.

### Judgments about Dates and Durations

Survey questions often ask about events that occurred within some specific time frame; for example, the National Crime Survey asks question (5):

(5) During the last 6 months, did anyone steal things that belonged to you from inside ANY car or truck, such as packages or clothing? [NCS]

Question (1b) is a similar item. Because people have difficulty remembering exact dates (Friedman, 1993; Thompson, Skowronski, Larsen, & Betz, 1996; Wagenaar, 1986), they may report events that took place before the specified *reference period* (i.e., the time period that the question asks about – the last six months in question (5)). This sort of error is known as *forward telescoping*, and this phenomenon has been the subject of theorizing by both survey methodologists and cognitive psychologists since Neter and Waksberg (1964) first documented it.

Reporting errors due to incorrect dating seem to arise through several distinct mechanisms. People may make incorrect inferences about timing based on the accessibility (or other properties) of the memory, incorrectly guess a date within an uncertain range, and round vague temporal information to prototypical values (such as 30 days). We take up these issues in detail in Chapter 4.

### Judgments about Frequencies

Another popular type of survey question asks about the number of times the respondent has engaged in some activity or about the rate of

Copyrighted material

events that happened in some reference period. Question (6) is typical of the sort of item we have in mind:

(6) During the past 12 months (that is since _(date)_ a year ago), about how many times did _____ see or talk to a medical doctor? (Do not count doctors seen while a patient in a hospital.) (Include the _____ visits you already told me about.) [HIS]

Individual visits to the doctor may be difficult to remember, but it is often possible for people to reconstruct what must have taken place (e.g., Means, Nigam, Zarrow, Loftus, & Donaldson, 1989; A. F. Smith, 1991; see also Lessler, Tourangeau, & Salter, 1989, on the reporting of dental visits). For example, people may recall only the regular pattern of events and use this rate to make an estimate for the full reference period (Burton & Blair, 1991). Chapter 5 considers the judgment processes involved in answering such questions.

### 1.2.4 Retrieval and Judgment for Attitude Questions

When we have to decide our attitude toward an issue, we need to consult our memory for relevant information. But what exactly do respondents retrieve when they answer a question about their attitudes? What does it mean for a respondent to give an inaccurate report about an attitude? Most attitude research seems to take it for granted that having an attitude means having a preexisting judgment about an issue or a person and that people automatically invoke these judgments when answering a relevant question (e.g., Fazio, 1989); responses to attitude questions do not, as a result, seem to present much of a problem or to require an elaborate cognitive theory.

We regard this view as overly simple. Along with Fischhoff (1991), we assume that there is a continuum corresponding to how well articulated a respondent's attitude is. At the more articulated end, the respondent has a preformed opinion just waiting to be offered to the interviewer; at the less articulated end, the respondent has no opinion whatever. Between these extremes, he or she may have a loosely related set of ideas to use in constructing an opinion or even a moderately well-formed viewpoint to draw on.

Respondents probably do not have perfectly well-articulated opinions on all attitude questions that survey researchers pose. First, evidence from large-scale attitude surveys indicates that on any given issue – no matter how familiar – a substantial portion of the population simply

Copyrighted material

does not have stable views. As Converse (1964, 1970) and others have shown, attitude responses over time sometimes show seemingly random shifts at the individual level even when no clear trends show up in the aggregate. Moreover, survey responses can shift dramatically in response to minor changes in question wording or order. Second, even if respondents do have more crystallized views about an issue, these views may not lend themselves to a clear-cut answer to the question at hand. The survey item may ask about an aspect of the issue that the respondent has not thought about. For instance, an item on the GSS asks whether abortions should be permitted in the case of rape; this item may give even ardent pro-life advocates reason to stop and reflect before they answer.

The judgment processes outlined in (4) are also relevant to attitude questions. For example, people may moderate or withhold their judgment if they feel that the information they possess is not sufficient (Yzerbyt, Schadron, Leyens, & Rocher, 1994); they may base their attitudes on what's most easily brought to mind (Ross & Sicoly, 1979); they may use stereotypes and schemas to fill in information they can't recall (e.g., Hastie, 1981); and they may combine piecemeal evaluations into a single assessment (Anderson, 1981). In Chapters 6 and 7, we present a detailed model of the attitude response process; we return there to the issue of when people base their answers on existing evaluations and when they base them on more specific considerations.

### 1.2.5 Reporting and Response Selection

The model's final set of component processes involves selecting and reporting an answer. We include two groups of processes here – mapping the answer onto the appropriate scale or response option and "editing" the response for consistency, acceptability, or other criteria.

Even when respondents have a clear answer to report, it may not be clear to them how to report it. The response options offered by the question are sometimes vague: In the case of attitude questions, where is the exact boundary between "Strongly agree" and "Agree"? In the domain of factual questions, how often do you have to eat in a restaurant in order for the frequency to qualify as "seldom" or "often"? Beyond the difficulties respondents may have with particular answer categories or response formats, they may differ in their approaches to selecting an answer. More than one response option may present a reasonable answer for a given respondent. Some respondents may work hard to choose

Copyrighted material

the best possible answer; others may be content to pick the first acceptable answer they consider. Respondents may also differ in their willingness to give an answer at all or to opt out of a question by saying they do not know the answer. Survey researchers have examined many questions regarding formats for answer categories, including how many categories should be offered, whether each one should be labeled, how the categories should be ordered, and whether "Don't know" should be included among them. We discuss these issues in more detail in Chapter 8.

Surveys often venture into areas that people do not ordinarily discuss with strangers. For example, a number of national surveys ask about the use of illicit drugs; other surveys ask about abortions, preferred methods of contraception, consumption of alcohol, medical conditions, or other topics that may cause embarrassment or resentment on the part of respondents. In fact, some respondents may regard one of the most basic of survey questions – *Who lives here?* – as an unwarranted intrusion (Tourangeau, Shapiro, Kearney, & Ernst, 1997). It is apparent that respondents do not always answer such questions truthfully. In Chapter 9, we consider the processes that govern the respondents' level of candor in answering sensitive questions – how respondents weigh the risks and benefits of responding truthfully (and of responding at all), what risks concern them, and what characteristics of the survey can raise or alleviate these concerns. Some respondents may decide in a rational way whether to respond truthfully or to give evasive answers; others may decide in an unconscious and automatic way, in accord with rules that originally evolved for dealing with other situations.

### 1.2.6 Summary

The theory of Table 1.1 sets out 13 cognitive processes that people may use to respond to a survey item. Despite the sheer number of processes, we should emphasize that survey responses are not the product of lengthy deliberations. On the contrary, respondents take less than 5 seconds to answer typical attitude questions (Bassili & Fletcher, 1991; Tourangeau, Rasinski, & D'Andrade, 1991).

In any model of this sort, a question immediately arises about the sequencing of the processes. Are the four main components distinct, nonoverlapping stages or are they simply classes of related processes? There is little evidence to help us decide this issue, and theoretical arguments can be made for either position. On the one hand, it seems likely

Copyrighted material

that for many questions respondents methodically follow the logical sequence of first comprehending the item, then retrieving the relevant facts, then making whatever judgment is called for, and finally reporting that judgment in the appropriate format. After all, how would respondents know what to retrieve before they have understood the question? How could they make a judgment if retrieval has not produced the necessary input? How could respondents report an answer before they have arrived at it?

On the other hand, it is equally clear that there can be many variations in the response process. In the first place, there will often be at least partial overlap among components. For example, retrieval is likely to commence before comprehension is complete; the very act of understanding key words in interpreting a question may trigger the spread of activation that is thought to be a central mechanism in retrieving memories (Anderson, 1983). In some cases, it may even make sense to think of retrieval as preceding comprehension; respondents may already have retrieved the information they need to answer one question in the process of answering earlier questions. Similarly, judgment may parallel rather than follow retrieval. In answering an attitude question, respondents may sequentially update their judgment about an issue as specific considerations come to mind. Or they may make judgments based on accessibility while they continue to retrieve memories. So any adequate model of the survey response process must allow for some overlapping processes.

In the second place, the model must allow respondents to backtrack from a "later" stage of the process to an "earlier" one. When retrieval yields little information, respondents may ask the interviewer for clarification or try to reinterpret the question on their own. At least one class of judgments involves deciding whether additional retrieval is warranted. An adequate model of the survey response process must therefore allow for cycling between the judgment and retrieval components. Similarly, selecting a response may require respondents to alter their judgment so that they can map the judgment onto one of the choices. Whenever the output from one component does not meet the requirement of another component, respondents may have to reexecute the earlier one.

A third complication is that respondents can sometimes skip or truncate a component. This accords with the cognitive toolbox approach that we mentioned in Section 1.2; respondents needn't employ all their response strategies in answering every question. Having understood a question as intrusive, respondents may skip retrieval and judgment alto-

Copyrighted material

gether and go directly to response selection – offering an evasive response or refusing to answer. In fact, inattentive respondents may even skip the comprehension stage and simply respond by saying they do not know the answer. With attitude items, respondents may omit the judgment component once they have retrieved an overall opinion (or at least a definite impression), provided that the form of the question permits an answer based on an existing judgment. With factual items, retrieval may yield an answer that requires little further cognitive processing; respondents do not need to use much in the way of judgment to answer questions about their age, date of birth, or sex. And, as we have already noted, judgment may replace retrieval when people answer on the basis of plausibility or familiarity rather than specific memories.

Because respondents can carry out components in parallel, because they can backtrack from later components to earlier ones, and because they can completely skip components, it would be misleading to describe the four components as nonoverlapping stages. Although we suspect that comprehension – retrieval – judgment – reporting is the most common order for the components, there is little evidence to support (or refute) this hypothesis, and other common patterns are worth noting.

## 1.3 Other Recent Proposals: High Road/Low Road Theories

The model in Table 1.1 constitutes an idealized list; the response process for a given item is likely to include only a subset of the processes identified there. Are there sequences other than the comprehension – retrieval – judgment – reporting sequence that are likely to arise in practice? As it happens, several models of the survey response process identify such sequences as alternative paths to responding.

We have already described the model proposed by Cannell and his colleagues in Section 1.1.2. This model features one track that includes most of the same processes that we have identified – comprehension, retrieval, judgment, and mapping – and a second track that modifies the response due to motives other than accuracy. This second track takes into account cues from the interview situation itself (see Figure 1.1).

We have also discussed some reasons for preferring a different model from that of Cannell and his colleagues. Here we consider an additional issue – the question of whether there are two distinct routes to arriving at an answer. It is certainly possible that, for some items, the conscientious and superficial routes represent sharply distinct tracks consisting of clearly distinguishable processes; however, in many cases the distinction between these two modes of processing is likely to blur. For example,

Copyrighted material

even when respondents are trying to be conscientious, prior items can affect their answers; indeed, the effects of retrieving information relevant to one item may have an automatic effect – an impact outside the respondent's awareness or control – on retrieval for related questions later on. Similarly, respondents may use both information retrieved from memory and cues from the situation in formulating their answers. There is no reason, in principle, why both sources could not contribute information to a response. Perhaps it is best to view Cannell's two tracks as two extremes on a continuum of processes that vary in the depth and the quality of thought that respondents give to their answers.

### 1.3.1 The Satisficing Model

Two other models of the survey response process have appeared in the last few years, and they share with the Cannell model the assumption of dual paths to a survey response – a high road and a low road. The first is the satisficing model that Krosnick and Alwin have presented (Krosnick & Alwin, 1987; see also Krosnick, 1991). Krosnick and Alwin distinguish between respondents who *satisfice* (the low road) and those who *optimize* (the high road) in answering survey questions. In their view, satisficing is not so much a strategy for choosing among response options as an overall approach to answering the questions (cf. Tourangeau, 1984). Satisficing respondents do not seek to understand the question completely, but just well enough to provide a plausible answer; they do not try to recall everything that is relevant, but just enough material on which to base an answer; and so on. Satisficing thus resembles the more superficial branch of Cannell's two-track model. Similarly, optimizing respondents would seem to follow the more careful branch.

Like the Cannell model, Krosnick's satisficing theory makes a sharp distinction among processes that probably vary continuously. Respondents may process questions to differing depths, and they may not carry out each component process with the same level of care. There is no reason to assume that respondents who are inattentive while reading a question will necessarily do a poor job of retrieval. The key point is that respondents may carry out each cognitive operation carefully or sloppily.

### 1.3.2 Strack and Martin's Two-Track Theory

Strack and Martin (1987) have also proposed a two-track model, one that focuses on the response process for attitude questions. The two

Copyrighted material

routes they identify correspond to the distinction between responses based on an existing judgment and those based on a new judgment that respondents derive at the time they answer the question. The route for new judgments closely parallels the model presented here; the key processes comprising that route are interpreting the question, accessing relevant information, "computing" the judgment, and formatting and editing the response. The other route leaves out the judgment step and replaces retrieval of more specific information with retrieval of a prior judgment. Although we certainly agree that processes resembling both tracks occur in surveys, we believe that both of Strack and Martin's tracks can be seen as special cases of the more general model we presented in Section 1.2.

There are several reasons for the choice of the general model over Strack and Martin's more specific one. First, we believe there is little evidence that respondents retrieve either an existing judgment or more specific beliefs about an issue but never retrieve both. In fact, in Chapter 6, we review evidence that a mix of specific beliefs and existing judgments may be the most common output of the retrieval process for attitude questions (in line with Fischhoff's, 1991, *partial perspectives* philosophy). In any case, it seems unnecessary to exclude this possibility a priori. Second, we question the model's assumption that, when people retrieve an earlier opinion, they circumvent the judgment process entirely. Although it may be possible to recall an attitude (e.g., that *The Last of the Mohicans* is boring) without being able to remember anything else about the topic, it is unclear how often this happens in surveys. Many attitude items require respondents to make new judgments even when they retrieve a prior opinion. For example, agree/disagree items such as those in (2) force respondents not only to retrieve their own views, but also to determine how close those views are to the position expressed by the item. Respondents must still perform some sort of comparison process, however abbreviated.

Like the routes in Figure 1.1, Strack and Martin's two branches represent pure cases. These ideal types are certainly worth noting, but it would be a mistake to see them as exhausting the possibilities.[4]

[4] Closer to our own approach is Schwarz's (1990) model for factual questions. Although this theory differentiates reports based on individual memories from those based on summary estimates (a high road/low road difference), it allows some variety in the relations between judgment and retrieval processes.

Copyrighted material

### 1.3.3 Summary

We prefer to think that quite a large number of paths to an answer are possible, depending on the effort that respondents are willing to invest and on the interplay between retrieval and judgment. In each case, the path traverses a subset of the processes identified here – which may be carried out well or sloppily, in parallel or in sequence, and with or without backtracking – as circumstances and motivation dictate.

A recent review by Jobe and Herrmann (1996) describes seven models of the response process, including several we have already discussed. Some of the others encompass the same processes listed in Table 1.1, differing from the present model only in how they group these processes into larger components. For example, Forsyth, Lessler, and Hubbard (1992) include separate comprehension and interpretation components; Willis, Royston, and Bercini (1991) include separate judgment and decision processes. These differences seem more a matter of emphasis than of substance. Two of the models attempt to account not only for the mental operations of the respondents but for those of the interviewer as well (Esposito & Jobe, 1991; Sander, Conrad, Mullin, & Herrmann, 1992).

## 1.4 Applications of the Model

Models like the ones summarized here are useful in part because they offer a new understanding of the sources of response effects in surveys and suggest methods for reducing such effects. For example, survey items such as (6) ask about the frequency of a specific class of events – visits to the doctor, days of illness, incidents of crime victimization, sexual partners, and so on. The outcome of each component of the response process will affect the number of events reported in answer to such questions. Depending on how respondents understand the question – in particular, on how broadly or narrowly they define the class of events to be reported – more or fewer events will qualify. Similarly, the number of events they report will reflect the outcomes of retrieval and judgment: The more relevant events they can remember, the greater the number they will report, and the more events they see as falling within the reference period for the question, the more events will be included in the answers.

Reporting errors are not the only place where cognitive models can shed light on survey problems. Chapters 7 and 10 of the book extend

Copyrighted material

this model to two additional problems encountered in surveys – the impact of the order of the questions and the effects of new computer-assisted methods of data collection.

One of the most puzzling sets of findings in the survey literature concerns the impact of item context. In Chapter 7, we describe our theory of how earlier questions in the interview or questionnaire can affect each component of the response process (Strack & Martin, 1987; Tourangeau & Rasinski, 1988). Prior items can change how respondents interpret later questions, what considerations they retrieve in formulating their answers, which standards or norms they apply in judging the issue, and how they report their answers. In addition, different mechanisms that produce context effects influence responses in different ways. Prior items sometimes influence later responses in the direction of consistency but sometimes have the opposite effect, producing apparent inconsistencies. Complicating matters further, the context of an item can change the overall direction of the answers or it can alter the correlations between answers to different items.

Over the last 25 years, the face of survey research has changed dramatically as computers have supplanted pencils and clipboards as the survey interviewer's most indispensable tools. Allowing the computer to collect the data has had a variety of effects, some of which have been subjected only to cursory investigation so far. The different methods of collecting survey data – self-administered questionnaires, interviews conducted by telephone or face-to-face, computer-assisted telephone and personal interviews, automated self-administered questionnaires – differ along many dimensions. We propose an analysis that singles out three key characteristics of the method of data collection – the degree of impersonality it conveys, the perception of legitimacy it fosters, and the level of cognitive burden it imposes on the respondent (and interviewer). Chapter 10 reviews the evidence regarding the effects of the method of data collection on the answers obtained.

## 1.5 Implications of the Model

In discussing the relationship between philosophy and psychology, William James once remarked that "metaphysics . . . spoils two good things when she injects herself into a natural science." There are those who might argue that the attempt to apply concepts and methods from the cognitive sciences to issues of survey methodology has had equally unhappy results. Innovations based on this attempt, sometimes dubbed the

Copyrighted material

*CASM movement* (for cognitive aspects of survey methodology), have been widely accepted within the federal statistical community, but there is also widespread scepticism about the value of this approach. In our final chapter, we consider the results that CASM has achieved so far.

Any effort to apply findings from one field to the problems of another raises several questions: Are researchers addressing the right problems? Are they applying valid theories? Have they developed the right applications of the theories for the problems at hand? We consider these questions in detail in Chapter 11, attempting to evaluate both our specific model and the more general movement to use cognitive science to improve survey practice.

Although the emphasis in this book is on the application of cognitive models to problems in survey methods, the findings in this growing literature clearly have something to offer the cognitive sciences in return. We single out several areas where survey findings have implications for cognitive theory. First, apart from survey-related studies, there have been few formal investigations of memory for everyday events, the happenings that are the stuff of daily life. Surveys routinely ask about consumer purchases, visits to the doctor, searches for jobs, hours spent at work, illnesses, hospital stays, courses taken in school, and a host of other daily occurrences. To be sure, a few landmark investigations of everyday memory have appeared within cognitive psychology (see Chapter 3), but these are meager compared to the large number of studies conducted with a view to improving survey reports about such issues.

The survey-inspired studies on everyday memory have yielded dividends bearing on several theoretical issues in the cognitive sciences. For example, some of the clearest evidence for the existence of generic memories – idealized memories of a class of recurring events – comes from attempts to understand survey reporting on dietary intake (A. F. Smith, 1991) and visits to the doctor (Means et al., 1989).

Similarly, investigations of how respondents answer questions about the frequency of everyday events have yielded rich insights into estimation processes and their role in filling in information that memory alone cannot provide. Herrmann (1992) describes some 15 strategies that respondents use in answering survey frequency questions, many of them involving estimation. To cite one example, Burton and Blair (1991) explore a strategy in which respondents recall events during a recent period and then make a rate-based projection for the entire period in question. This strategy is far closer to the estimation procedures employed by statisticians or engineers than to the heuristics that have taken

Copyrighted material

center stage in discussions of frequency judgments within the psychological literature. Investigations of rounding in survey reporting (e.g., Huttenlocher, Hedges, & Bradburn, 1990) have shed further light on how respondents compensate for vague or incomplete memories. Since the time of Bartlett's classic work (1932), psychologists have acknowledged that memory involves both retrieval and reconstruction; the survey-based studies on generic memory, estimation, and rounding have added considerable detail to this picture.

Another area where the survey literature has much to offer cognitive psychology involves what might be called *proxy memory* – memories for events experienced by other people. This topic has been almost completely neglected within the mainstream memory literature (see Larsen, 1988, for an exception) but has been a lively area within the movement to apply cognitive theories to issues of survey methods (e.g., Blair, Menon, & Bickart, 1991).

In addition to its implications for the study of memory, the recent efforts to apply cognitive theories to survey issues have much to contribute to the study of attitudes. As we've remarked, investigations of attitudes in social psychology often seem to assume that respondents have a preexisting answer to most attitude questions and need only to read out this answer. The results from the survey literature present quite a different picture: Responses to survey questions can become unreliable over time (Converse, 1964, 1970) and show fluctuations as a consequence of seemingly minor changes in question wording (Schuman & Presser, 1981). In fact, simply changing the order of the questions can produce large swings in the answers (Tourangeau & Rasinski, 1988). If answers to attitude questions are simply readouts of stored judgments, it is not clear why question order should make such a difference. The study of order effects on responses to attitude questions has been a particularly fruitful area for the application of cognitive methods to a long-standing survey problem (Schwarz & Sudman, 1992; Tourangeau & Rasinski, 1988). These models are aimed at explaining survey results, but they have greatly expanded our understanding of assimilation and contrast effects in judgment more generally.

In several areas, then, the effort to apply concepts and methods drawn from psychology to problems in surveys has yielded benefits to both fields. Still, the sailing hasn't always been smooth. In Chapter 11, we consider some of the barriers to further progress.

Copyrighted material