CHAPTER

3

Remembering What Happened: Memory Errors and Survey Reports

Roger Tourangeau
The Gallup Organization, Rockville, MD

A central tool of social science research—perhaps *the* central tool—is asking people questions about what happened. Because of the critical role of retrospective reports, a major source of error in social science data is memory error. This chapter presents a brief overview of memory and its contribution to error in self-reports. It begins by examining models of autobiographical memory and then explores the processes responsible for forgetting and other memory errors.

AUTOBIOGRAPHICAL MEMORY

Perhaps the most elementary question we can ask about memory is what gets remembered. What sort of thing do we store in memory? It is immediately apparent that are at least three distinct types of material in memory—facts culled from books or oral descriptions, personal experiences, and knowledge about how to do things. What most of us think of as memory consists largely of memories for personal experiences, and it is this sort of memory—autobiographical memory—that is usually at issue when we gather self-report data. There is general agreement that the basic format in which events or personal experiences are encoded is that of the story. We experience our lives as organized around actors who have intentions and carry out plans that succeed or fail; these mini-narratives are the stuff of which autobiographical memory is composed.

Memory Structure

If the unit of autobiographical memory is the experience or event, then the next issue is how they are represented and structured. There are two basic models of the structure of autobiographical memory. According to one, memories form an associative network (e.g., Anderson, 1983; see Collins & Quillian, 1969, for an early version); according to the other, memories form a hierarchy (Barsalou, 1988; Kolodner, 1985). Network models represent concepts or ideas as nodes in the network, and the relationship between concepts as links. These links reflect basic semantic relations between concepts such as the subset–superset relation. Network models typically assume that we search memory by tracing the links between concepts. Retrieval is the process of activating a concept—bringing it to consciousness—and activation is thought to spread from one concept to another along the links of the memory network.

According to the hierarchical model of memory, similar events are organized into categories. The categories may encompass both subtypes and individual experiences. For example, in Kolodner's (1985) hierarchical model, the category-level representations are called event memory organization packages (EMOPs, for short). New EMOPs are formed when a sufficient number of memories—several events of the same type—are stored under an existing EMOP. Memory search begins with the relevant category (the EMOP) and proceeds downward as indices (distinguishing particulars) are generated that define ever finer categories until individual experiences are found. The Kolodner model is an example of a schema-plus-tag theory, in which memories are seen as consisting of a general pattern stored at the category level (the schema) plus one or more individuating details (the tags). The general pattern for a class of events (say, doctor visits) might include information about the usual participants (doctors, nurses, patients, receptionists, and so on), the typical location (the doctor's office or HMO), the larger sequence to which the event might be linked (the treatment of a chronic condition), the superordinate category to which this one belongs, and the subcategories of this class of experiences (visits to different types of doctors).

These ideas—that a memory for an experience includes both generic and unique information, that retrieval encompasses both automatic and controlled processes (e.g., spreading activation and the generation of retrieval cues), and that memory search consists of generating progressively more specific cues—are widely shared, even by memory researchers who do not subscribe to other assumptions of the Anderson or Kolodner models.

Sources of Forgetting

Within this framework, we can distinguish at least four major classes of memory problems. The first involves encoding. We may never form a representation of an event in the first place or the representation that we do form may be so sketchy as to render retrieval difficult or impossible. When there are serious encoding problems, little or no information ever reaches long-term memory. A second class of problems involves errors introduced *after* the original encoding; things that happen after the original experience can be woven into the representation of the experience and distort our memory of it. Another class of memory problems involves retrieval failure. We may have encoded the experience adequately in the first place and have preserved this representation intact, but when it is time to recall the event, we may simply be unable to remember it. A final class of errors involves reconstruction. When we fill in details missing from the memory, based on our general knowledge of a class of events or our expectations about change over time, we may introduce inaccuracies. In some cases, retrospective reports are not based on specific recollections at all, but represent estimates or other types of inferences (Bradburn, Rips, & Shevell, 1987; Burton & Blair, 1991; Menon & Yorkston, chapter 5, this volume).

ENCODING PROBLEMS

The most extreme encoding problems involve cases in which the relevant information is never noticed in the first place and so never enters memory, but there can be less extreme encoding problems as well. To the extent that the initial encoding of the information is superficial, the information will be hard to remember later on. The principle that deeper encoding leads to better recall was initially established for passages of text (Craik & Lockhart, 1972); subjects remembered a passage better when their task required them to process it deeply (e.g., to summarize the passage) than when it required them to process it only superficially (to count the number of words in it). Later work suggests that the more elaborated the initial representation, the more likely the experience can be retrieved later on (Anderson & Reder, 1978). According to a network model, elaborative encoding establishes multiple links to the experience, each of which can serve as a path for retrieving it; similarly, in a hierarchical model, elaborative encoding produces multiple indices that can be used to search memory for the experience.

Aside from the problems resulting from insufficient or superficial encoding of an experience, the initial representation of an event can give rise to a second type of memory problem—it may fail to match the retrieval cue provided later on (Tulving & Thomson, 1973). When the experience is stored in one category (or EMOP) but the retrieval cue triggers the search of a different one, we are unlikely to recall the experience. In fact, as Tulving and Thomson demonstrated, if the retrieval cue makes us think of the wrong category, we may not even recognize the relevant experience.

We recently carried out a study in which encoding problems appeared to be the main culprit in producing reporting errors in a survey (Brittingham, Lee, Tourangeau, & Willis, in press). The study examined parents' reports about their children's vaccinations. Two national surveys monitor rates of vaccination coverage among children. Both surveys ask parents about their children's vaccinations, encouraging them to consult the "shot cards" that many pediatricians provide; these cards record information about each vaccination the child has received. Unfortunately, many parents do not have the cards and are forced to rely on unaided recall instead. Such reports are known to be error-prone (e.g., Goldstein, Kviz, & Daum, 1993). Our study investigated whether parents take in the information about which shots their children got in the first place. Parents may dutifully take their children for their scheduled shots without paying too much attention to which ones were administered on any specific occasion.

We interviewed a sample of parents as they were leaving an HMO, just after their child had received one or more vaccinations. The questionnaire asked parents to describe what happened during the visit and probed them specifically about any immunizations the child received. Table 3.1 shows the main results from the study, displaying several measures of the accuracy of the parents' reports. The results can be summarized very simply: Even as the parents were leaving the doctor's office, their reports were close to chance levels of accuracy. The overall correlations between the reports and the records were significant for only three of the five immunizations and

TABLE 3.1
Immunization Study Results

	Accuracy Measure			
Vaccine	*False Negative Rate*	*False Positive Rate*	Phi	*Net Bias*
Hepatitis B	51.7%	20.0%	.20	−41.4
	(60)	(10)	(70)	(70)
DTP (Diphtheria-Tetanus-Pertussis)	41.4%	16.7%	.32*	−31.4
	(58)	(12)	(70)	(70)
Polio	33.9%	14.3%	.42*	−24.3
	(56)	(14)	(70)	(70)
Hib (Haemophilus Influenzae b)	86.5%	0.0%	.20	−64.3
	(52)	(18)	(70)	(70)
MMR (Mumps-Measles-Rubella)	33.3%	19.4%	.23*	17.1
	(3)	(67)	(70)	(70)

Note. Parenthetical entries are cell sizes; asterisks indicate a *phi* correlation significant at $p < .05$. The false negative rate is the percentage of parents who failed to report a vaccination the child received that day. The false positive rate is percentage who reported a vaccination the child had not received that day. *Phi* is the overall correlation between the parent's report and the clinic's records about whether the child had received a given vaccine. Net bias refers to the difference between the percentage of parents reporting the vaccination and the percentage of children actually receiving it.

Source: Brittingham et al. (in press).

these three correlations were none too impressive. We also conducted a follow-up interview with parents 10 weeks later. Performance after 10 weeks was not much worse than it was after a few minutes. If they took in the information at all, the parents were, for most part, able remember it over the 10-week interval.

Why was accuracy so low? As a general rule, the depth and elaboration of the encoding of an event reflects such variables as its distinctiveness, emotional impact, and duration. Unusual or dramatic events, or those that unfold over a long period of time, tend to grab our attention and hold it long enough to ensure that a rich representation is created and stored in long-term memory. Childhood vaccinations have none of these characteristics; to the contrary, they are frequent, routine, and quick. When we did our study, children were supposed to have received at least 14 doses of five different vaccines by their second birthday. (Since then, a sixth vaccine has been added to the recommended list.) The long and technical names of the vaccines (Haemophilus influenzae b), the relative unfamiliarity of the illnesses they prevent, and the administration of multiple vaccines during a single visit may also inhibit accurate encoding.

So, one potential source of error in retrospective reports is that the respondent never really knew the answer in the first place. We may fail to encode enough information to produce an accurate account of the experience later—even when later is only a few minutes after the event took place.

STORAGE PROBLEMS: THE INCORPORATION OF POST-EVENT INFORMATION

A second class of memory problems arises after the initial encoding of the event; it involves what happens to the memory while it is being stored in long-term memory. Rehearsal—time spent thinking or talking about the event—is thought to play a key role in maintaining the accessibility of a memory. Take flashbulb memories as an example. These are the peculiarly vivid memories left by events like the assassination of President Kennedy or the explosion of the space shuttle *Challenger* (Brown & Kulik, 1977; see also Conway, 1995). The literature on such memories singles out two major groups of variables believed to affect their level of detail—the amount of rehearsal and the degree of surprise or other emotion initially engendered by the event (e.g., Pillemer, 1984; Rubin & Kozin, 1984; Winograd & Killinger, 1983). The level of emotional impact probably affects the elaboration of the encoding of the experience, but rehearsal probably affects its long-term accessibility to retrieval. Although some studies of flashbulb memories indicate that the importance of rehearsal may have been overstated, the consensus is that rehearsal plays a key role in maintaining detailed and vivid memories over long spans of time.

Of course, as Neisser and Harsch (1992) demonstrated, the presence of a detailed memory is no guarantee of its accuracy. In fact, the process of recounting an event may not only preserve the memory but add details to it, often inaccurate ones (e.g., Loftus & Kaufman, 1992). Our recollection of an experience may change every time we recount it. Details of the event may be elaborated or abbreviated depending on the context in which the event is described, and any errors introduced in the telling may become part of the memory for the event. Rehearsal may produce greater error when the memory is vague or nonexistent to begin with. Perhaps no researcher has demonstrated the difficulties we face in distinguishing what we actually experienced from what we heard, thought, or learned later on than Loftus (see chapter 12, this volume). These difficulties derive from our tendency to incorporate "postevent" information into our representation of the event without distinguishing its source. What we experienced firsthand may differ in vividness or detail from what we said or heard later on, but the differences between the two types of information may fade as time passes.

The same mechanisms that account for the distorting effects of postevent information (including self-generated postevent information) may also explain another common type of memory error—reporting something that did not happen. In studies of memory, such errors are referred to as false alarms or intrusion errors. Once again, the problem is in distinguishing what actually happened from what sounds good, seems to fit, or was merely imagined. These reporting errors are all the more likely when memories for what did occur are indistinct, reducing the difference between events that were actually experienced and those that were merely heard about or imagined. Two studies reported by Johnson, Foley, Suengas, and Raye (1988) illustrate these problems; the studies found very few differences between subjects' memories for actual childhood events and their memories for dreams and fantasies from childhood. By contrast, there were many differences between memories for actual and imagined adult experiences. By adding detail and increasing the apparent familiarity of an event, repeated rehearsal or visualization can further reduce the difference between memories for actual and imagined events.

Three factors affect whether a memory is accepted as genuine (i.e., based on an actual experience rather than a fantasy or secondhand report)—the qualities of the memory itself, its overall plausibility, and the strictness of the standard used in judging the memory's genuineness.

The Qualities of the Memory

Memories seen as arising from direct experience differ from those seen as originating in other sources (e.g., reading or imagination) in several ways. Memories judged to be based on direct experience include more perceptual detail than those thought to be derived from other sources (Johnson, Hashtroudi, & Lindsay, 1993); similarly, results from the flashbulb memory literature suggest that the vividness of the memory (presumably reflecting the presence of perceptual details) is related to confidence in the accuracy of the memory (Neisser & Harsch, 1992). Memories judged to be based on actual experience are also more likely to include peripheral details of the event and less likely to include information about cognitive processes (Johnson et al., 1993; Schooler, Gerhard, & Loftus, 1986). A final variable that may affect judgments about the source of a memory is the ease of retrieving it (Jacoby, Kelley, & Dywan, 1989). Unfortunately, none of these variables are infallible guides to the source of a memory. Repeated rehearsal can affect ease of retrieval; visualization can add perceptual detail. Techniques used to make recall easier may simultaneously make it harder to distinguish real from imagined incidents (cf. Lindsay & Read, 1994).

Relative Plausibility

At least some of the time, judgments about the accuracy of a memory reflect information beyond that contained in the memory itself. The presence of conflicting (or corroborating) evidence clearly affects judgments about whether a memory is genuine. For example, the literature on eyewitness accuracy indicates that delay between the event and the recall attempt can affect the acceptance of misleading postevent information (Lindsay, 1990; Loftus & Hoffman, 1989), a finding that may reflect the diminished accessibility of conflicting information in memory derived from the original experience. Supporting evidence from other witnesses can also increase acceptance of false postevent information, even of information that implicates oneself as the guilty party (Kassin & Kiechel, 1996).

Decision Criteria

Some tasks may require only a relatively low threshold for deciding to accept a memory; others may demand a more careful sifting of the evidence. It is one thing to claim to recognize someone who seems to know us, quite another to identify a potential criminal in a lineup (cf. Lindsay & Johnson, 1989). External pressures (such as fatigue or social pressure) or characteristics of the person making the judgment (such as his or her youth or suggestibility) may bias the judgment as to whether the memory is real (Johnson, Kounios, & Reeder, 1994).

In short, memory is not judgment-free. What we retrieve from memory often consists of our current beliefs about an incident, beliefs that reflect what we actually experienced (and remember), what we did not experience but infer, and what we learned later on. The problem is that it can be difficult to distinguish between beliefs acquired through direct experience and those acquired through other means.

RETRIEVAL FAILURE

Retrieval failure is often cited as the most common source of forgetting. It occurs when information is stored in long-term memory but we are unable to get it out. Problems with retrieval loom large as a source of forgetting because it is clear that memory often contains far more information than we think it does. Unaided or free recall is likely to yield fewer memories than recall aided by retrieval cues or hints, and cued recall generally yields fewer memories than recognition. The more help we give the retrieval process, the more information it seems to turn up, including some things we thought were quite beyond its reach. In addition, since Ebbinghaus' (1885) initial explorations of the phenomenon, it has been known that even when we seem to have completely forgotten a topic, we often demonstrate savings in learning the material over again. Something has been retained in memory that makes it easier to acquire the information a second time than to learn it initially. Finally, even severely amnesic patients may show implicit memory; for example, they are more likely to complete a word missing one or more letters (HO_S_) with a word they had seen earlier (HORSE), even though they cannot recall the earlier word itself (Warrington & Weiskrantz, 1968). These patients stored *something* in memory even if it was beyond the reach of retrieval and could no longer be made conscious. (See Schacter, 1987, for a more thorough discussion of implicit memory). All of these findings suggest that a major source of forgetting is failure to retrieve information that is still there.

One of the most obvious facts about forgetting is its relation to the passage of time. No single variable seems to have such a profound impact on the accessibility of a memory than its age. Most theories of memory attribute this loss of accessibility over time to the interfering effects of later experiences. The problem is that the characteristics that made the experience unique initially are shared with later experiences; as a result, the original event may get lost among the similar events experienced afterwards. It is easy to recall our only trip to a doctor; it is far more difficult to pick out a particular trip when we have made dozens of similar ones.

Both the network and hierarchical models of the structure of long-term memory offer ready accounts for the interfering effects of later experiences. Both types of models assume that when similar events are experienced, a "generic" memory is formed, which leaves out the details of the individual incidents but records their overall pattern. These generic memories—EMOPs in Kolodner's model—explain why it so much easier to recall what usually happens than to recall the specific details that distinguish one incident of a given type from another (see also Means & Loftus, 1991; and Smith, Jobe, & Mingay, 1991).

According to the network model, multiple incidents of the same type (such as doctor visits) are likely to be linked to a single node (the node representing our concept of doctor visits). As we try to recall an individual event by thinking about the general concept, activation spreads along the links leading from that concept to the nodes representing individual experiences. Unfortunately, the more links leading away from the node where retrieval begins, the more that activation is dissipated across these links and the less likely that the memory we seek will receive enough activation to be retrieved (see Fig. 3.1). Anderson (1983) referred to this as the "fan effect." According to the hierarchical model, as we experience similar events, we form new subcategories (in the Kolodner model, new EMOPs) that capture finer distinctions among the events. The creation of these subcategories can impose an added burden on the retrieval process, requiring ever more detailed indices to be generated in order to locate the specific memory.

Retention Curves

The effects of the passage of time on recall accuracy have been demonstrated with almost every kind of event (e.g., Rubin & Wetzel, 1996). For example, a recent review of the survey literature found reduced levels of reporting or reduced reporting accuracy for hospital stays, health care visits, medical conditions, dietary intake, smoking, car accidents, hunting and

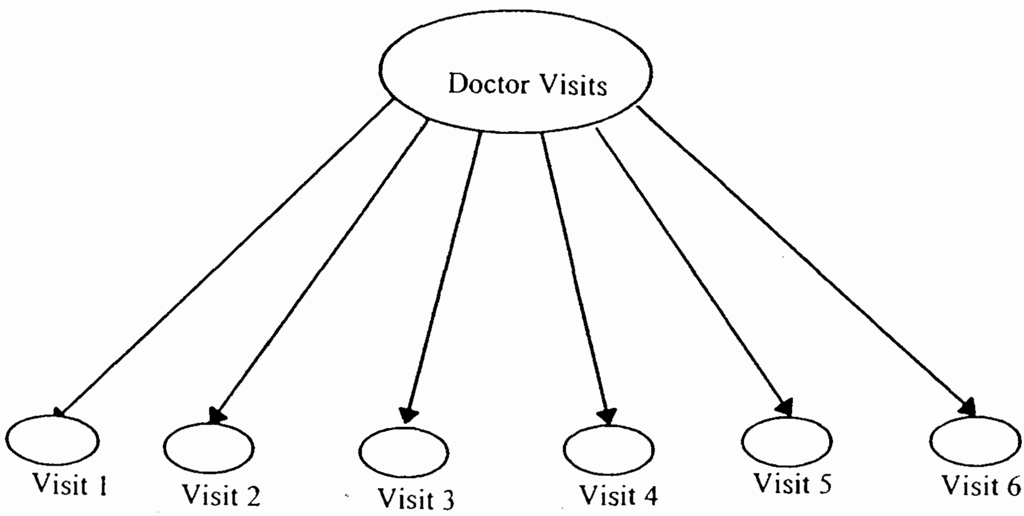

FIG. 3.1. Network representation of the general concept of doctor visits and of memories for individual visits. Activation may not reach the nodes representing the individual visits because of the large number of links leading away from the node for the general concept of doctor visits.

fishing trips, consumer purchases, and home repairs as the length of the retention interval increased (Jobe, Tourangeau, & Smith, 1993). Given the importance of these effects, it is hardly surprising that there have been a number of attempts to find the empirical function that best relates the amount of information retained in memory to the length of the retention interval. At least four functional forms have been proposed (Rubin & Wetzel, 1996).

Exponential Decay

Suppose we retain a fixed proportion of the information we learned originally over each unit of time that passes. For example, we may remember 90% of the items we purchased a month ago. After 2 months, we remember 81% of them (90% of 90%), and so on. If p denotes the fixed proportion retained each month, then after t months we will remember p^t. Under this model, the proportion retained (r) after t units of time have passed is:

$$\begin{aligned} r &= ap^t \\ &= ae^{bt} \end{aligned}$$

in which a reflects the level of initial learning (the percent retained at time zero) and b is the natural logarithm of p (cf. Sudman & Bradburn, 1973). The parameter a is useful for situations (like the immunization study) in which the initial encoding of the memory is not perfect.

Hyperbolic Decay

If the factor responsible for the decline of memory over time is the accumulation of similar events, then a different functional form may capture the impact of the passage of time more accurately. Suppose events accumulate at a rate of b events per unit of time; for example, we might purchase three or four items during the average month. The proportion of events retained in memory over a time period of length t would be inversely related to the total number of similar events that occurred over that period:

$$r = \frac{1}{a + bt},$$

in which a again reflects performance at time zero.

Logarithmic Decay

A third functional form has been suggested, based on the idea that equal ratios of elapsed time should produce equal amounts of memory loss. As Rubin and Wetzel (1996) observed:

> There is an easy way to arrive at the logarithmic function if in psychological terms equal ratios of time, not equal intervals, are important. Assume that the psychological difference between the 3 to 4 ratio of 3 and 4 seconds is the same psychological difference as that between 30 and 40 minutes, or 18 and 24 hr, or 3 and 4 decades . . . The simplest function to describe retention is the linear function, $y = -m \cdot x + b$. If one uses the logarithm of time, as suggested by the equal ratios observation, instead of time for x, this equation becomes the logarithmic equation. (p. 749)

Expressing the logarithmic equation in the same format as the exponential and hyperbolic functions yields:

$$r = a - b \ln(t).$$

It is sometimes easier to work with $\ln(t + 1)$, which again makes a the level of initial performance.

Power Function

The final family of curves is based on the power function:

$$r = \frac{a}{(t+1)^b}.$$

(Again, $t + 1$ is used in the denominator instead of t for the sake of mathematical convenience.) This model implies equal ratios of retention with equal ratios of time (i.e., the ratio between the proportion retained at 5 and 10 years will be the same as the ratio between retention at 1 and 2 days). The power function has received some empirical support from work by Anderson and Schooler (1991), and Rubin and Wetzel's (1996) meta-analysis suggests that, of the four functions, it provides the best fit to the autobiographical memory data.

Still, all four functional forms share several basic predictions—that forgetting increases monotonically over time, but that it occurs rapidly at first and then slows down. In addition, it is not too difficult to find values for b so that the different models yield similar quantitative predictions. In Fig. 3.2, the value of a has been set to 1 for all four models (that is, performance is perfect at the outset) and values of b were found that yielded nearly identical predictions for Time 1. As the figure illustrates, the shape of the four curves is quite similar and the divergence among them is not very noticeable until a relatively long time has passed. It can, then, be quite difficult to distinguish the different models empirically (but see Bradburn et al., 1987, for a somewhat different view on this issue).

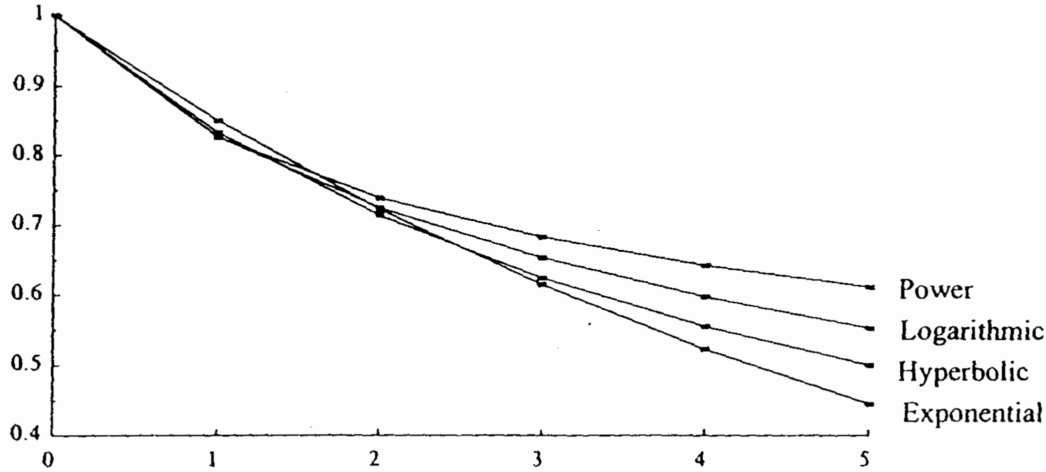

FIG. 3.2. Theoretical retention curves. The four curves demonstrate the similar shapes of the four functions when the key parameters are deliberately set to produce similar predictions.

Aiding Retrieval

Researchers have investigated a number of methods for improving recall. For example, it is apparent that simply taking more time to remember can help. Although there is doubtless some upper bound on this strategy, in one study (Williams & Hollan, 1981), subjects continued to recall new items even after nine previous sessions, each lasting an hour. Even a little extra time can produce benefits (Cannell, Miller, & Oksenberg, 1981).

Other methods thought to increase the amount of information retrieved include:

- Decomposition, a strategy that breaks a class of events down into subclasses, each of which is then recalled separately;
- Recalling events in reverse chronological order (that is, beginning with the most recent event and working backwards in time), a strategy that seems to help sometimes (Loftus & Fathi, 1985; Loftus, Smith, Klinger, & Fiedler, 1992) but not always (Jobe et al., 1990; Smith, 1991);
- Listing temporal boundaries or landmarks, such as major life transitions, to aid in the recall of events near those landmark events (Means & Loftus, 1991).

The decomposition strategy converts a free recall task into a cued recall task with the subclasses serving as additional retrieval cues. For example, it is probably easier to recall all of our doctor visits if we are cued to think about each of the major specialties. The findings about the usefulness of temporal boundaries (such as moving to a new city or starting a job) suggest that autobiographical memories are organized not only into categories based on the type of event, but also by periods in a person's life. Other results also indicate that we think of our lives as organized into major phases or chapters bounded by temporal landmarks–graduations, weddings, new jobs. Events near these landmarks seem to be remembered more easily than those occurring farther from the boundaries between periods (e.g., Robinson, 1986).

Number and Types of Cues

We do not remember every aspect of an event equally well. It is often difficult to remember names, and exact dates are notoriously difficult to remember. By contrast, we recognize faces quite readily. Several careful studies of autobiographical memory confirm these everyday observations. Perhaps the best of these is the monumental study by Wagenaar (1986), who recorded some 2,400 events from his daily life over the course of 6 years. Wagenaar noted what happened ("had dinner with friends"), who took part, and when and where the event took place. He tested his memory for each type of fact about the events later on. The results indicated that the different aspects of the events were not equally memorable. Regardless of the initial cue Wagenaar used to jog his memory, he had the most difficulty recalling the date of the event (to within a week of the actual date). On the other hand, there were no real differences among the levels of recall for the participants, activities, and locations of the events. In addition, the different aspects of the events were not equally effective as retrieval cues; information about the nature of the event was the best cue for retrieving other facts about the event and information about when an event occurred was the worst cue.

The effectiveness of a given type of cue will depend on several variables, including the nature of the memory task and the encoding of the event. If the task is to retrieve a particular event, then the best cue will be the one that most successfully distinguishes that event from similar events–given that the cue is consistent with the initial encoding of the event. In the extreme case, some representation of the event or information itself will be the best cue; it is, for example, far easier to recognize someone's name than to recall that name with the aid of some other cue. In Wagenaar's study, the description of what happened was probably the most effective at distinguishing the event in question; persons and places can be linked to many experiences and so are less helpful in picking out any specific one. Although time cues are potentially quite distinctive to the target event, it appears that we rarely encode our experiences with exact dates (see Bradburn, chapter 4, this volume); as a result, time cues do not help us very much. Different considerations apply when the task is to retrieve as many events as possible. In that case, cues linked to many incidents may have the advantage. Barsalou's (1988) results suggest that, with this task, locations may prompt the recall of more events than activities or times and that activities and times may outperform persons as retrieval cues.

RECONSTRUCTION ERRORS

As Wagenaar's (1986) results illustrate, retrieval often yields only partial results, turning up some details of an experience but not others. When this happens, we often attempt to fill in or reconstruct the missing pieces. Unfortunately, as Bartlett (1932) first demonstrated nearly 70 years ago, the details we add are not always an accurate representation of what actually happened.

A major source of the details added during this reconstruction process is our general knowledge about what is typical for events of a given type. Thus, the reconstructed event may deviate from the original in a systematic direction—that of resembling too closely the typical pattern for events of its type. Bartlett's subjects tended to drop particularly odd details from the Indian folk tales they were trying to remember and to add things that made the stories a little more sensible. In interpreting stories, we often supply the unstated connections between events (e.g., Bower, Black, & Turner, 1979); it seems clear that we do the same sort of filling in the gaps in memories for firsthand experiences as well. Because the details we add conform to the pattern for the situation, individual peculiarities are lost and the representation of the individual events comes to resemble the generic representation for events of that type. Inferences that fill in missing or implicit details may be made at the time of encoding, later on during subsequent rehearsal, or at the time of retrieval, but regardless of when they are made, the inferences are likely to be similar in content.

There may also be relatively subtle differences among the processes of generating retrieval cues that bring back memories of actual experiences, making reasonable inferences about details missing from partial memories, and imagining plausible scenarios that never really took place. In each case, we are likely to imagine what might have occurred based on our general knowledge. The result is sometimes an accurate recollection of a real experience, sometimes a plausible reconstruction of what might have occurred, and sometimes a complete fabrication. With distant or poorly recalled events, it may be difficult to tell these three situations apart.

Retrospective Biases

Numerous studies have examined situations in which respondents are asked to report on some personal characteristic—such as their views about an attitude issue—and then are asked at some later time to report both the current value of that characteristic and their earlier answers as well. The results often suggest that the current value is used as an anchor on which the "memory" of the past value is based. In an early demonstration, Bem and McConnell (1970) assessed subjects' attitudes, exposed them to a procedure that produced large changes in their attitudes, and then asked some of the subjects whose attitudes had changed to report their current views and others to recall their earlier reports about their attitudes. The group reporting their initial attitudes gave answers that closely paralleled those who reported their current views; in fact, as Bem and McConnell observed, "The figures are so similar . . . that it would appear that we had asked subjects for their current attitudes rather than their initial attitudes" (p. 28). The phenomenon has been replicated several times since Bem and McConnell's initial demonstration (e.g., Smith, 1984; see Ross, 1988, for a review of the findings). The impact of our current state on our recollection of the past is apparent for other types of memory as well—such as our recall of pain, past use of illicit substances, or income in the last year (see Pearson, Ross, & Dawes, 1992, for further examples).

These results suggest that we may reconstruct the past by consulting the present and projecting it backwards, assuming more stability in the characteristic or behavior in question than it actually exhibits. There are, on the other hand, times when we seem to exaggerate the amount of change we have undergone. Conway and Ross (1984) reported that persons who completed a self-help program rated their preprogram skills lower after they completed the program than they had beforehand; persons on a waiting list showed no comparable changes in recalling their initial level of skill. Believing they had changed, those who completed the program apparently exaggerated the amount of improvement they had experienced.

The exaggeration of both consistency and change may reflect a single underlying process—the attempt to reconstruct the past using the present as an anchor. We may adjust this anchor but we seem to underadjust when we expect the characteristic to be stable and to overadjust when we expect it to change. Either way, what we "recall" is, in fact, a kind of estimate.

Estimation Versus Recall

Other findings also suggest that inference and estimation processes can supplant retrieval as the basis for retrospective reports. Burton and Blair (1991) identified two main strategies for answering behavior frequency questions (e.g., questions about the number of doctor visits in the past 6 months). One strategy, which they call episode enumeration, is to recall and count individual incidents; the other, rate-based estimation, is to project the typical rate over the length of the recall period. Burton and Blair found that respondents were less likely to use episode enumeration as the period covered by the questions grew longer, as there were more episodes to recall, and as the respondents were more rushed for time. As it becomes more difficult to recall each incident, we reconstruct the total by estimating it. (See Menon & Yorkston, chapter 5, this volume, for additional findings on the role of estimation in answers to questions about behavioral frequencies.)

Estimation and inference are likely to have an impact with other memory tasks as well. For example, a study by Huttenlocher, Hedges, and Bradburn (1990) examined answers to a question about how long ago an event had taken place. The further back in time the event was, the more likely the answers were to be reported as round numbers (such as "40 days ago"). These results suggest that, as it became harder to remember exactly when the event occurred, respondents switched to estimation strategies that yielded only approximate answers. Similarly, in a recognition task, we may judge whether we encountered an item before based on its overall plausibility rather than its familiarity (Reder, 1987). Such inferential processes may be especially important in helping us to distinguish events we have experienced but forgotten from those we never experienced at all. Our judgment about whether we experienced the event may hinge on whether it is the sort of thing we think we would remember (Gentner & Collins, 1981).

CONCLUSIONS

Several processes can make it hard for us to remember what happened. We do not notice everything that happens; as a result, some experiences are never stored in memory at all and the ones that are stored may lack key details, such as information about exact dates. Once a representation of the experience does enter memory, it does not necessarily remain static over time. New information, including inferences we draw about the experience or embellishments we add in recounting it, can become part of the memory. Even when a memory has been preserved intact, it may be difficult or impossible to retrieve it. The accumulation of similar experiences over time seems to be the chief source of difficulties in retrieval, producing a rapid drop in our chances of retrieving the item over the short run and slower drops as further time passes (see Fig. 3.1). The best cues to help us recall an experience are ones that most clearly distinguish the experience from all the others with which it might be confused, but even a cue that uniquely picks out the experience will not trigger the retrieval of the memory if it does not match the event's representation in memory. We often try to fill in what we cannot retrieve, using inferential processes or estimation strategies based on our general assumptions about different types of events. These reconstruction processes are a final source of memory errors.

There can be a fine line between retrieving a memory and inferring what might have happened. Autobiographical memory seems designed to record our current beliefs about the past; one aspect of those beliefs that does not seem particularly well represented in memory is their source. As memories fade, we lose our ability to distinguish those parts of our picture of the past that were derived from direct experience from those added through inference, secondhand reports, or imagination. The fact that we can remember something–even have vivid and detailed memories for it–carries no guarantee that we remember it accurately.

REFERENCES

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R., & Reder, L. (1978). An elaborative processing explanation of depth of processing. In L. Cermak & F. I. M. Craik (Ed.), *Levels of processing and human memory* (pp. 385–403). Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2*, 396–408.

Barsalou, L. W. (1988). The content and organization of autobiographical memories. In U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 193–243). Cambridge, England: Cambridge University Press.

Bartlett, F. (1932). *Remembering–A study in experimental and social psychology*. Cambridge, England: Cambridge University Press.

Bem, D. J., & McConnell, H. K. (1970). Testing the self-perception explanation of dissonance phenomena: On the salience of premanipulation attitudes. *Journal of Personality and Social Psychology, 14*, 23–31.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology, 11*, 177–220.

Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science, 236*, 157–161.

Brittingham, A., Lee, L., Tourangeau, R., & Willis, G. (in press). Errors in parents' reports of children's immunizations. *Vital & Health Statistics*. Series 6.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5*, 73–99.

Burton, S., & Blair, E. (1991). Task conditions, response formulation processes, and response accuracy for behavioral frequency questions in surveys. *Public Opinion Quarterly, 55*, 50–79.

Cannell, C. F., Miller, P. V., & Oksenberg, L. (1981). Research on interviewing techniques. In S. Leinhardt (Ed.), *Sociological methodology 1981* (pp. 389–437). San Francisco: Jossey-Bass.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*, 240–247.

Conway, M. A. (1995). *Flashbulb memories*. Hove, UK: Lawrence Erlbaum Associates.

Conway, M. A., & Ross, M. (1984). Getting what you want by revising what you had. *Journal of Personality and Social Psychology, 47*, 738–748.

Craik, F. I. M., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.

Ebbinghaus, H. (1885). *Uber das Gedächtnis*. Leipzig: Duncker and Humblot.

Gentner, D., & Collins, A. (1981). Studies of inference from lack of knowledge. *Memory and Cognition, 9*, 434–443.

Goldstein, K. P., Kviz, F. J., & Daum, R. S. (1993). Accuracy of immunization histories provided by adults accompanying preschool children to a pediatric emergency department. *Journal of the American Medical Association, 270*, 2190–2194.

Huttenlocher, J., Hedges, L., & Bradburn, N. M. (1990). Reports of elapsed time: Bounding and rounding processes in estimation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 196–213.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 391–422). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jobe, J. B., Tourangeau, R., & Smith, A. F. (1993). Contributions of survey research to the understanding of memory. *Applied Cognitive Psychology, 7*, 567–584.

Jobe, J. B., White, A. A., Kelley, C. L., Mingay, D. J., Sanchez, M. J., & Loftus, E. F. (1990). Recall strategies and memory for health care visits. *The Milbank Quarterly, 68*, 171–189.

Johnson, M. K., Foley, M., Suengas, A., & Raye, C. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General, 117*, 371–376.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 144*, 3–28.

Johnson, M. K., Kounios, J., & Reeder, J. A. (1994). Time-course studies of reality monitoring and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1409–1419.

Kassin, S. M., & Kiechel, K. L. (1996). The social psychology of false confessions: Compliance, internalization, and confabulation. *Psychological Science, 7*, 125–128.

Kolodner, J. (1985). Memory for experience. In G. H. Bower (Ed.), *The psychology of learning and motivation* (vol. 19, pp. 1–57). Orlando, FL: Academic Press.

Lindsay, D. S. (1990). Misleading suggestions can impair eyewitnesses' ability to remember event details. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 1077–1083.

Lindsay, D. S., & Johnson, M. (1989). The eyewitness suggestibility effect and memory for source. *Memory and Cognition, 17*, 349–358.

Lindsay, D. S., & Read, J. D. (1994). Psychotherapy and memories of childhood sexual abuse: A cognitive perspective. *Applied Cognitive Psychology, 8*, 281–338.

Loftus, E. F., & Fathi, D. C. (1985). Retrieving multiple autobiographical memories. *Social Cognition, 3*, 280–295.

Loftus, E. F., & Hoffman, H. G. (1989). Misinformation and memory: The creation of new memories. *Journal of Experimental Psychology: General, 118*, 100–104.

Loftus, E. F., & Kaufman, L. (1992). Why do traumatic experiences sometimes produce good memory (flashbulbs) and sometimes no memory (repression)? In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of "flashbulb" memories* (pp. 212–223). New York: Cambridge University Press.

Loftus, E. F., Smith, K. D., Klinger, M. R., & Fiedler, J. (1992). Memory and mismemory for health events. In J. M. Tanur (Ed.), *Questions about questions: Inquiries into the cognitive bases of surveys* (pp. 102–137). New York: Sage.

Means, B., & Loftus, E. (1991). When personal history repeats itself: Decomposing memories for recurring events. *Applied Cognitive Psychology, 5*, 297–318.

Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about *Challenger*. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of "flashbulb" memories* (pp. 9–31). Cambridge, England: Cambridge University Press.

Pearson, R. W., Ross, M., & Dawes, R. M. (1992). Personal recall and the limits of retrospective questions in surveys. J. M. Tanur (Ed.), *Questions about questions: Inquiries into the cognitive basis of surveys* (pp. 65–94). New York: Sage.

Pillemer, D. B. (1984). Flashbulb memories of the assassination attempt on President Reagan *Cognition, 16*, 63–80.

Reder, L. (1987). Strategy selection in question answering. *Cognitive Psychology, 19*, 90–138.

Robinson, J. A. (1986). Temporal reference systems and autobiographical memory. In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 159–188). Cambridge, England: Cambridge University Press.

Ross, M. (1988). The relation of implicit theories to the construction of personal histories. *Psychological Review*, *96*, 341–357.

Rubin, D. C., & Kozin, M. (1984). Vivid memories. *Cognition*, *16*, 81–95.

Rubin, D. C., & Wetzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, *103*, 734–760.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 501–518.

Schooler, J. W., Gerhard, D., & Loftus, E. F. (1986). Qualities of the unreal. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 171–181.

Smith, A. F. (1991). Cognitive processes in long-term dietary recall. *Vital and Health Statistics*, Series 6, No. 4 (DHHS Publication No. PHS 92-1079). Washington, DC: U.S. Government Printing Office.

Smith, A. F., Jobe, J., & Mingay, D. (1991). Retrieval from memory of dietary information. *Applied Cognitive Psychology*, *5*, 269–296.

Smith, T. W. (1984). Recalling attitudes: An analysis of retrospective questions on the 1982 General Social Survey. *Public Opinion Quarterly*, *48*, 639–649.

Sudman, S., & Bradburn, N. M. (1973). Effects of time and memory factors on response in surveys. *Journal of the American Statistical Association*, *68*, 805–815.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352–373.

Wagenaar, W. (1986). My memory: A study of autobiographical memory over six years. *Cognitive Psychology*, *18*, 225–252.

Warrington, E. K., & Weiskrantz, L. (1968). New method of testing long-term retention with special reference to amnesic patients. *Nature*, *217*, 972–974.

Williams, M. D., & Hollan, J. D. (1981). The process of retrieval from very long-term memory. *Cognitive Science*, *5*, 87–119.

Winograd, E., & Killinger, W. A. (1983). Relating age at encoding in early childhood to adult recall: Development of flashbulb memories. *Journal of Experimental Psychology: General*, *112*, 413–422.

Lawrence Erlbaum Associates, Inc., Publishers
10 Industrial Avenue
Mahwah, NJ 07430

Cover design by Kathryn Houghtaling Lacey

Library of Congress Cataloging-in-Publication Data

The science of self-report : implications for research and practice / edited by Arthur A. Stone . . . [et al.].

p. cm.

Includes bibliographical references and index.
ISBN 0-8058-2990-3 (cloth : alk. paper).—ISBN 0-8058-2991-1 (pbk : alk. paper)
Human experimentation in medicine–Congresses. 2. Human experimentation on psychology–Congresses. I. Stone, Arthur A.
R853.H8S36 1999
619—dc21 98-33155
CIP

Books published by Lawrence Erlbaum Associates are printed on acid-free paper, and their bindings are chosen for strength and durability.

Printed in the United States of America
10 9 8 7 6 5 4 3 2