The ecological study of memory

ULRIC NEISSER

Department of Psychology, Cornell University, Ithaca, NY 14853, USA

SUMMARY

The study of memory has long been dominated by the structural tradition, and especially by the experimental analysis of mechanisms of information processing. That dominance may soon be brought to an end by the progress of neuroscience, which offers more direct ways of studying the mechanisms in question. At that point functional issues may move to centre stage. Those issues include the act of remembering and its social functions, the skills and presuppositions of the rememberer, the interaction of those skills and presuppositions with the particular material being remembered, and the determinants of accuracy and confabulation in recall.

1. INTRODUCTION

There are two fundamentally different ways to think about memory. Correspondingly, there have long been two distinct streams of memory research. Both are flourishing today; both have a bright future. The distinction between the two streams is not primarily one of method (say, between laboratory studies and field studies), nor is it just based on differences between disciplines involved (say, between neuroscience and psychology). More fundamental than either of those, it is the distinction between the study of structure and of function in memory. One research tradition focuses on the storage of information from the past, and on the neural structures that make such storage possible. The other, in contrast, focuses on how stored information is used in the present. This distinction is by no means new: readers familiar with the history of the study of memory will recognize echoes of Ebbinghaus and Bartlett in my argument.

The first of these research traditions, the one focused primarily on structure, has always been the more popular and productive. Hundreds of psychologists have published thousands of papers on structural aspects of memory, and the rate of publication is increasing even as we speak. According to Koriat & Goldsmith (1996) this approach reflects the classical metaphor of a storehouse. In the storehouse metaphor, people are said to 'put information' into memory at one time and 'take it out' again at another, just as valuable objects can be put into storage and later retrieved. This familiar metaphor has generated many useful questions over the years: how much information can the brain's storehouse hold? Could there be several of them? If so, do they have different properties and different capacities? Until recently, such questions were largely metaphorical. Today, however, neuroscience is moving beyond metaphor to uncover the real mechanisms responsible for information storage in the brain. Instead of a passive storehouse, there is a dynamic complex of systems that interact in ways we are only now beginning to understand.

Other papers in this volume expound on those systems. My own interest is in the second stream of memory research, i.e. in the functional tradition. Here the basic metaphor is not storage but action. The act of remembering typically has a purpose: it is a goal-directed action that someone undertakes for a particular reason at a particular time and place. Memories are not so much retrieved as they are constructed, usually with a specific aim in mind. Because it is goal-directed, remembering can be regarded as more or less successful. If the goal happens to be fidelity to some original event, it can also be regarded as more or less accurate. But quite apart from questions of accuracy, we can always ask what it was about this setting, this material, and this individual that led to the production of just this memory in just this form. These are good questions too, and quite as difficult as those generated by the storehouse metaphor. I think of them as ecological questions, because they concern the relations between rememberers and environments. The rest of this paper describes some recent and interesting ecological answers.

2. CONSTRAINTS ON RECALL

Frederick Bartlett (1932) was the first psychologist to describe remembering as 'constructive'. He meant that ordinary recall is almost never the exact reproduction of something old, of stored stimuli or responses in their original forms. Instead it is typically the production of something new, appropriate to the remember's present situation. Bartlett's famous studies of story recall illustrate this point. The subjects began by reading a story and trying to recall it as accurately as they could. Later, often much later, they were asked to recall it

again, not just once but a number of times. Bartlett found that the recollected stories changed from one recall to the next, often undergoing significant shifts of meaning as the subjects tried to make sense out of whatever information was still available to them. To obtain similar distortions in a shorter period of time, Bartlett often used what he called 'serial recall'. In this familiar parlour game, A tells the story to B, who tells it to C, who tells it to D. Serial recall makes an excellent class demonstration, often producing hilarious errors after only a few steps in the chain.

Because Bartlettian distortion has often been described in textbooks, such changes may no longer surprise us. What is surprising, in contrast, is that some materials seem to be immune to them. Consider, as a first example, the simple counting-out rhymes that children chant as they decide who is to be 'it':

Eenie, meenie, miney, mo;
Catch a tiger by the toe;
If he hollers, let him go;
Eenie, meenie, miney, mo.

Scholars have been collecting these rhymes for over a century; it turns out that they persist almost unchanged from one generation of children to the next. How can this happen, given that they are transmitted exclusively through serial recall?

To move from the childish to the sublime, let us consider another example: the oral epic poems of the Balkans. This was the genre of Homer, and until recently it was still practised by skilled, and entirely illiterate, singers in the mountains of former Yugoslavia. Such a performer can sing all evening, night after night: he knows many different songs. Each song consists of thousands of 'lines', lines that have been passed down orally from one singer to the next for centuries. Why haven't they long since been distorted or even destroyed by the constructive processes of memory?

Fortunately these are not mere rhetorical questions; we now know most of the answers. The recent work of David Rubin, as documented in his book *Memory in oral traditions* (1995), explains exactly why these materials are so memorable: they are subject to multiple constraints. In the case of 'Eenie, meenie, miney, mo' and other counting-out rhymes, the constraints are imposed by such poetic devices as rhyme, alliteration, metre and front-to-back vowel progression. (The progression in 'Eeenie, meenie, miney, mo' is very similar to that in 'Fee, fie, foe, fum'.) Given these poetic restrictions, very few one-word changes in 'Eenie, meenie, miney, mo' are even possible. When change does occur (as it recently has for tiger), it happens at only a few predictable points in the text and draws from only a predictable set of alternative words.

A different set of constraints operates in classical oral poetry; the Balkan epics include very little alliteration or rhyme. Nevertheless, every word in the poem as performed must fit both the meaning and the metre. The poets tailor their songs to the metrical constraints, as they sing, drawing on a familiar repertoire of formulaic expressions to make things come out right at the end of each line. In the *Odyssey*, for example, a small stock of epithets for each character appears over and over: 'noble' Odysseus, 'brilliant' Odysseus, 'long-suffering' Odysseus, sometimes 'long-suffering brilliant' Odysseus. Which of them appears in a given line depends on metrical considerations. These constraints are so firm that they can preserve the essential core of an epic through centuries of oral transmission.

Are constraints really so powerful? To see how rapidly a wide range of alternatives can be narrowed to a single choice, consider this example suggested by Rubin (1995). If I ask you to think of a colour, you have a great many options. If I were to ask you to think of a word that rhymes with 'bed', there would again be many possibilities. But if you try to think of a colour that rhymes with bed, you will find exactly one and no more. That is the situation that often prevails in poetry and song. In such cases, remembering is still a constructive process, but the construction is so tightly constrained that it produces the same result almost every time.

3. IMPLICIT THEORIES

Personal and autobiographical memories are also typically constrained, but in a different way. Michael Ross (1989) has described these constraints as 'implicit theories'. Consider, for example, two recent studies in which clinical patients were asked to remember the intensity of their own pains. In one study, Eich *et al.* (1985) asked chronic headache sufferers to make hourly ratings of their pain during the day. Then, on arrival at the clinic, they (i) rated their degree of pain at that moment and (ii) tried to recall the levels of pain they had experienced since their last visit. The results were clear. Patients who happened to be experiencing strong pain at the time of report tended to overestimate past pain levels, whereas those experiencing little pain tended to underestimate those levels. These patients apparently had an implicit theory that their headaches would be much the same from day to day, and used that theory systematically in estimating their earlier pain levels. In their case, what seemed to be simple retrieval was actually theoretical inference.

The situation is quite different for patients who have had some form of treatment. In a different study, Linton & Melin (1982) asked individuals who were about to undergo therapy for chronic pain to make baseline ratings of their pain intensity. After the programme was over, the same persons tried to remember what those initial ratings had been. These recalls exhibited systematic underestimation: 11 of 12 patients remembered their baseline pain as higher than it had really been. They too had an implicit theory: namely, that treatment helps. If I hurt this much now, after all that therapy, my pain at baseline must have been even worse!

Ross (1989) provides many examples of such inferences. My favourite involves college students who applied to a study-skill clinic. First, they rated their baseline skill levels. Then half of them were actually enrolled in the programme; the others were assigned to a waiting list. At the end of the clinic, both groups

tried to remember the baseline ratings that they had given earlier. The wait-listed (control) group showed no bias, but most subjects in the participant group remembered their initial skill ratings as lower than they had actually been. Like the patients in the pain clinic, they assumed that the treatment must have done them some good. These examples show that remembering is not just a matter of retrieving stored facts; it is a purposeful action that individuals support with whatever information is available to them.

4. THE SOCIAL CONTEXT

Ross's (1989) subjects, like the participants in most other memory experiments, were trying to give accurate reports about past events. In more natural settings, however, accuracy is rarely the main goal of recall. A recent study by Ira Hyman (1994) vividly illustrates this point. Hyman varied both the context and the purpose of remembering. His subjects (all Emory undergraduates) began by reading a de Maupassant short story. Some then reported individually to the experimenter, as is usual in memory research. Others met instead with a peer, i.e. another subject who had also read the same story. Both groups were further subdivided by their instructions: half of them were to 'remember' what they had read, while the other half were asked to say 'what they got out of the story'.

Both social context and instructions turned out to be important. Subjects who talked with their peers provided less detailed narratives than those who talked to experimenters, but made many more evaluative and reflective comments. They often described personal reactions: perhaps they liked the story, or felt sorry for one of the characters, or were struck by the customs of 19th-century France. Those who met with the experimenter were much less likely to say such things. There were similar contrasts between subjects instructed to remember and those asked to say what the story had meant to them. Looking across the whole four-cell design, the group most closely modelled on standard laboratory procedures ('Tell the experimenter what you remember') was maximally different from the group in the most natural situation ('Talk with this other student about what the story meant to you'). By focusing exclusively on directed recall, the standard methods of cognitive science may miss many of the normal uses of declarative memory.

The subjects of Hyman's experiment probably knew that they were adapting their behaviour to the situation. Given instructions for systematic recall, they recalled systematically; given a situation that seemed to call for conversation, they conversed. But memory is not always so easily controlled, and people do not always understand the basis of their own performance. Even the most confident rememberers can be wrong about their own recollections: memories that seem to be direct and unmediated may instead be based on sheer narrative construction. To illustrate this point I will conclude with two of my own studies, both focused on vivid memories of particular events. Although the studies were very similar in design, they produced opposite results. The contrast between them may help us to understand yet another ecological constraint on the process of remembering.

5. FLASHBULB MEMORIES

The phrase 'flashbulb memory' was coined in a specific context: American informants who were trying to recall how they had first heard about the 1963 assassination of John F. Kennedy. Twelve years after the event, almost all the subjects interviewed by Brown & Kulik (1977) said they still remembered that moment very clearly. Taking these reports at face value, Brown & Kulik ascribed their accuracy to a special, virtually infallible memory system that comes into play only at moments of great stress and surprise. Later investigators, including myself (Neisser 1982), have been sceptical of this claim: after all, there was no way to determine whether the recollections were accurate.

To help resolve this issue, I undertook a new study (Neisser & Harsch 1992). When the space shuttle *Challenger* exploded in January of 1986, the nationwide shock was so great that I thought it might produce a new set of 'flashbulb memories'. The next morning, in a colleague's freshman classroom, I distributed a questionnaire on which students recorded how they had first learned about the event on the previous day: who had first told them about it, where they were, what they were doing, who else was there, what time it was. I put the completed questionnaires away until the fall of 1989, nearly three years later, when the erstwhile freshmen had become seniors. Two students working with me, Jeff Gutkin and Nicole Harsch, contacted those who were still at Emory and asked them to come to the lab for an experiment. There they were given a questionnaire just like the one they had filled out three years earlier, except that now they also rated their confidence in each aspect of the memory. Thus, unlike Brown & Kulik (1977), we had a way of checking the accuracy of the memories themselves.

The results were surprising. I myself would have been satisfied to find occasional small errors, enough to show that such memories are not infallible. But many of the errors were not small: instead, the subjects were dead wrong. One of them, who confidently remembered learning about the disaster from TV when an announcer broke in with the news, had in fact heard about it from fellow students in her Religion class. Another, who was quite sure he had been at home breakfasting with his family, had in fact been on campus. A woman student gave a dramatic account of a girl running through the dorm screaming 'The space shuttle blew up'; in fact, she had heard about it from friends over lunch. Of course not everyone was mistaken to this extent; many reports were partially right and a few almost completely so. But the means score on our seven-point accuracy scale was only 2.95, and a quarter of the subjects were at zero. Nevertheless, almost all of them were very sure they were right. Harsch later interviewed most of these subjects personally. They all stuck to their stories despite fairly strong hints that they might be wrong. At the end of the

interview, each of them was shown his or her original 1986 questionnaire. Most of them were astonished: they recognized their handwriting and admitted that the earlier account 'must be right', but insisted that they 'still remember it this other way'. Mistaken memories may be very permanent, and they can be invested with seriously misplaced confidence.

Where did all those errors come from? Some of them, especially those that reflect what we called 'TV priority', are easy to understand. This term refers to cases where subjects falsely believed that they first heard about the disaster from television when in fact they had encountered it in some other, more personal way. TV priority is easily explained by Ross's (1989) concept of implicit theories: many people believe that one most often learns of disasters from television. Such errors may also be examples of what Brewer (1988) calls 'wrong time slice' memories: the recalled event really did happen, but is not the right answer to the experimenter's question. It is probable that all of our subjects, however they may have first learned of the disaster, ended up watching it on television later that day. The dramatic explosion images that were shown so repeatedly on every TV network may have stuck in their minds, until finally they came to believe that those images had been their own first exposure to the news of the explosion.

These hypotheses are plausible, but many of the errors in our study cannot be so easily explained. What about the entirely fictitious girl who ran screaming through the dorm, for example? My own guess is that the subject who produced this vivid memory had herself been very upset by the explosion, so upset that she felt like screaming. Later, she 'projected' those unuttered screams on to an imagined other. Is that how it really was? In individual cases, we can never know.

At a more general level, I believe that many errors in confident memories result from the following three-step process:

1. First, the subject has to forget what really happened.
2. Then, as the individual develops a new narrative to fill in the gaps, factors like implicit theories and wrong time slices and projective fantasies come into play.
3. Finally, that reconstructed narrative becomes so familiar that the subject accepts it as valid. If this hypothesis is right, the process of false memory creation typically requires an initial forgetting phase. Only when some amount of forgetting has occurred do other mechanisms come into play. This suggests a further question: what would happen if there was so much immediate rehearsal that individuals had no opportunity to forget? The next study sheds some light on this point.

6. THE EARTHQUAKE STUDY

One limitation of our space shuttle experiment (and of many similar studies) is that the subjects were not personally involved in the action. Individuals may be startled or upset or dismayed to learn of some distant calamity, but still nothing much is happening to them. Would they remember better if they were personally involved? To be sure, 'being there' does not guarantee recall in all cases: we eat lunch personally every day but soon forget most of our lunches. But, however its effect may have to be qualified, 'personal involvement' is certainly an intriguing variable. To study its effects more closely, my colleagues and I took advantage of another disaster.

On 17 October 1989, the Loma Prieta earthquake shook San Francisco and other communities in northern California. It started fires, collapsed highways, interrupted the World Series, broke the Bay Bridge. Early the following day I called Steve Palmer, a friend and cognitive psychologist in Berkeley, and we roughed out plans for a study of earthquake memories. The project eventually grew to include informants in three different cities: 41 subjects in Berkeley, where the effects of the quake were rather mild; 44 in Santa Cruz, where it was more severe; and 76 in Atlanta, where it was just another news event. The project researchers included Gene Winograd and Erik Bergman in Atlanta, Palmer and Charles Schreiber in Berkeley, and Mary Sue Weldon in Santa Cruz.

All informants filled out questionnaires a few days after the earthquake. Both California groups, in Berkeley and in Santa Cruz, were asked to describe their experiences during the quake and to rate their emotional reactions. The Atlanta subjects reported how they had heard the news. For technical reasons we were unable to wait three years as in the shuttle study: recall was tested after a year and a half.

A full account of the results appears elsewhere (Neisser *et al*. 1996); here there is only space for the bottom line. I was surprised again. Unlike the Atlanta controls (and also unlike the subjects of the *Challenger* study), our Berkeley and Santa Cruz informants remembered their experiences almost perfectly. Their recall of the moment when the earthquake hit couldn't have been better: they forgot almost nothing and got nothing wrong. Even their recall of less personally involving matters ('How did you hear the news about the Bay Bridge?') was substantially better than that of controls at Emory. How can we explain this finding? Why were the earthquake memories so much more accurate than the space shuttle memories had been?

Given the poor performance of our Atlanta controls, the shorter recall interval cannot have been responsible. In some sense the key factor must have the 'personal involvement' of the California subjects, but how did their 'involvement' lead to more accurate recall? Given recent work on the physiological effects of emotional arousal (see LeDoux & Muller, this volume), the most obvious hypothesis would attribute the gain to emotional arousal. More specifically, one might suggest that earthquakes frighten people, and that the resulting physiological processes act to strengthen the memory trace.

Perhaps surprisingly, several aspects of our data speak against this hypothesis. For one thing, there were no significant correlations between rated arousal and amount of recall. (This was not just due to ceiling effects: correlations were negligible even for items that had lower overall levels of accuracy, like 'How did you

hear about the collapse of the Bay Bridge?') More importantly, the mean rated arousal of the California subjects was just not very high. Indeed, there was no reason why it should have been. The actual impact in Berkeley was no stronger than that of other tremors our Californian subjects had experienced many times before; there was little reason for alarm. Three Berkeley subjects did not notice the quake at all, but a year and a half later all of them remembered just what they were doing while not noticing it!

If the stress of the moment was not what made these experiences so memorable, the key factor must have been something that came into play afterward. In my view, that something made its appearance as soon as people realized that this had been no ordinary tremor. They had been in a 'big one': the Bay Bridge was broken, San Francisco was burning. What did they do then? They started to talk about it. People who live through 'big' events find many opportunities to describe their experiences, not just once, but over and over. They tell their stories to everyone they meet, and also to everyone who calls on the phone to see if they are all right. My hypothesis is that those rehearsals, the telling and retelling of each individual's earthquake narrative, are what made the experience of this particular earthquake especially memorable.

Earthquake narratives were so common and continuous in California in late 1989 that they eventually became a joke: people began to wear T-shirts that said 'Thank you for not sharing your earthquake experiences'. The effect of these retellings must have been to strengthen the tellers' own memories of their experiences before they had a chance to forget what had happened. As a result, the error-producing processes discussed earlier in this chapter had no space in which to exert their effects. Memory depends not only on structures and mechanisms in the brain but on critical aspects of the local ecology: here, on the narrative demands of the social situation.

The future of both research traditions in the study of memory is bright indeed. In the coming years we will surely learn a great deal about the neural systems that preserve information in the brain. In addition, we can expect to learn a lot about the act of remembering itself: how it depends on the constraints of the material, on the social situation, on what the rememberer is trying to do, on what happens between acquisition and recall. Both of these approaches are necessary if we are ever to understand the exquisitely human activity of remembering in an adequate way.

REFERENCES

Bartlett, F. C. 1932 *Remembering*. Cambridge University Press.

Brewer, W. F. 1988 Memory for randomly sampled autobiographical events. In *Remembering reconsidered: ecological and traditional approaches to the study of memory* (ed. U. Neisser & E. Winograd), pp. 21–90. New York: Cambridge University Press.

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

Eich, E., Reeves, J. L., Jaeger, B. & Graff-Redford, S. B. 1985 Memory for pain: relation between past and present pain intensity. *Pain* **23**, 75–380.

Hyman, I. E. 1994 Conversational remembering: story recall with a peer versus for an experimenter. *Appl. Cogn. Psychol.* **8**, 4–66.

Koriat, A. & Goldsmith, M. 1996 Memory metaphors and the real life/laboratory controversy: correspondence versus storehouse conceptions of memory. *Behav. Brain Sci.* **19**, 167–228.

Linton, S. J. & Melin, L. 1982 The accuracy of remembering chronic pain. *Pain* **13**, 281–285.

Neisser, U. 1982 Snapshots or benchmarks. In *Memory observed: remembering in natural contexts* (ed. U. Neisser), pp. 43–48. New York: Freeman.

Neisser, U. & Harsch, N. 1992 Phantom flashbulbs: false recollections of hearing the news about Challenger. In *Affect and accuracy in recall: studies of flashbulb memories* (ed. E. Winograd & U. Neisser), pp. 9–31. New York: Cambridge University Press.

Neisser, U., Winograd, E., Bergman, E. T., Schreiber, C. A., Palmer, S. E. & Weldon, M. S. 1996 Remembering the earthquake: direct experience vs. hearing the news. *Memory* **4** 337–357.

Ross, M. 1989 Relation of implicit theories to the construction of personal histories. *Psychol. Rev.* **96**, 341–357.

Rubin, D. C. 1995 *Memory in oral traditions: the cognitive psychology of epic, ballads, and counting-out rhymes*. New York: Oxford University Press.