# Paradoxes of learning and memory

Henry L. Roediger, III and Andrew C. Butler

## Summary

We explore 12 paradoxes of learning, memory and knowing in our chapter. These are mysteries in which subjective experience – what we think we know or remember – does not correspond to objective facts. In some cases, we hold false memories: we are utterly confident in our memories that events happened one way, but they did not. Another example is hindsight bias: we may believe that we knew (after the fact) how an event would turn out, but controlled experiments show people cannot predict the event. Another category of illusion occurs with learning. Often students judge one method of learning to be superior to a second method, but their actual performance shows the reverse to be true. The paradox of interference creates other puzzles: when people try to remember similar events, they will often confuse one for another. We discuss 12 paradoxes and their implications for cognitive functioning. Some of these errors may implicate cognitive strategies that we use because they often lead to correct answers in many situations, but can produce errors in other instances.

## Introduction

Psychologists love mysteries and paradoxes. They always have, they always will. There is nothing surprising here; all people like paradoxes and puzzles. Look at Figure 8.1 and ask yourself which surface of the two boxes is longer, the one on the left or the one on the right? Every person naïve to the situation will answer the one on the left. However, the two surfaces are exactly congruent. They are the same. Try tracing over one and laying it over the other if you don't believe us. Even if you are aware that it is an illusion (created by Shepard, 1981), you still fall for it every time. Knowing the two surfaces are the same does not correct our perception of the boxes. Psychologists studying perception have discovered hundreds of remarkable illusions like this one and people generally find them fascinating.

Cognitive and social psychologists also love illusions. Our journals are filled with puzzles of the following sort: clever experimenters manipulate a variable that has a large effect on some judgement or behaviour, then they ask subjects to predict their own behaviour in the situation and show that, lo and behold, the subjects either give random predictions or make completely wrong ones. For example, variable A increases a behaviour whereas the subject thinks the variable decreased it or had no effect. A variant on this theme is to show how human behaviour violates the rules of some normative theory about behaviour. The field of behavioural economics has grown up around observations that
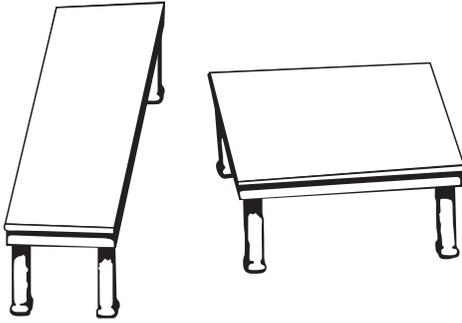
**Figure 8.1** A spatial illusion created by Roger Shepard (1981). The surface of the left box appears much wider than the one on the right, but the two are actually congruent.

people fail to follow the rational models of 'economic man' in making decisions about money. The predictions work neither on the microeconomic scale of individual human behaviour (e.g. Kahneman, 2003) nor the macroeconomic scale of national and international finance (e.g. witness the world's economies thrown into complete disarray in 2007–2008, which virtually no economist predicted).

When subjects in our experiments are asked to explain their behaviour, they often make up a coherent story, even if it is one that is wildly inaccurate and does not account for the facts. Nisbett and Wilson (1977) reviewed many studies from social and cognitive psychology and argued that people 'tell more than they can know'. If we clearly do not know the real causes of the behaviour (the independent variable that the experimenter manipulated), we make up a good enough story, nonetheless. In a later book, Wilson (2002) argued that these tendencies at self-delusion are so pervasive that we are 'strangers to ourselves' (the title of his book). Books by Dunning (2005) and Gilovich (1991) make similar points.

Psychologists are much less successful at explaining the mysteries they raise. We report interesting puzzles, we explore them experimentally, we root around in them for a while, and then we move on to the next puzzle (as has been pointed out by critics; e.g. Newell, 1973). This may be an unfortunate tradition, but it is one we generally follow in writing this chapter. We write about 12 interesting paradoxes that have been uncovered using behavioural paradigms (our chapter is a 'no-brainer'). Yes, we know that the title of the book involves the brain, but we will have nothing whatsoever to say about the neural bases of the illusions and puzzles that we review. We can be confident that the brain holds the secrets to all these phenomena, but neural explanations for them are not at hand (but for a few potential leads, see the section on *Future Challenges and Questions* at the end of the chapter).

We organize our chapter into four main parts. The first section is concerned with paradoxes of remembering and knowing. Why do people suffer false memories, remembering some event differently from the way it happened, or remembering an event that never happened at all? Often these erroneous memories are held with high confidence. Or why does our knowledge sometimes blind us to the way others see the world? An expert has trouble seeing the world through the eyes of the novice, even though the expert was once a novice (see Chapter 9). The second section of the chapter is concerned with paradoxes of task difficulty and students' judgements of their own learning and memory as a function of difficulty. The surprising finding is that people often misjudge the conditions of learning that lead to good retention later. Even students, who are expert learners, have such erroneous beliefs.

A third section is concerned with paradoxes of interference. The basic feature of these puzzles is that when people try to remember an event that happened some time ago, they can become confused by events that happened more recently, during the intervening time since the original event. The more recent events can interfere with memory for the original event that the person is trying to retrieve. A final section of the chapter deals with puzzles caused by the fluency with which cognitive processes are carried out. Our being able to perceive or remember something easily colours the weight we give that information in making judgements of the world. We often overweight information that is easy to perceive or retrieve.

We cover these illusions and paradoxes as though they are separate, but they probably have some common causes. We pause along the way to make connections where appropriate. Our chapter is perforce rather superficial – we identify a paradox, puzzle or illusion, and then we move on to the next. Many articles, chapters and even books have been written about these phenomena. Following the citations in each section via Google or some other search engine would bring about a wealth of information. Pohl's (2004) edited volume is a good place to find more information about many of these phenomena.

## Paradoxes of remembering and knowing

We all believe in our memories, the record of our lives. Memories contain our identity; to believe that cherished memories might be false could mean that our self-image is wrong, too. Yet even strongly held memories can turn out to be wrong. Jean Piaget, a pioneer in the study of cognitive development, had an early memory of this sort. A critical moment in his life occurred when his nanny was walking him in a carriage on a street in Paris. A kidnapper tried to steal him, but his nanny fought back and saved him. Piaget later wrote: 'I was held in by the strap fastened round me while my nurse bravely tried to stand between me and the thief. She received various scratches, and I can still vaguely see those on her face'. However, when he was 15, Piaget's parents received a letter from the nurse, who had recently been converted and was confessing past sins. She had made up the whole story, faking the scratches, and she returned a valuable watch she had been given for her bravery in the situation. Piaget remarked 'I therefore must have heard, as a child, the account of this story . . . and projected it into the past in the form of a visual memory'. He further opined that 'Many real memories are doubtless of the same order' (quotes are from Piaget, 1962, pp. 187–8).

Of course, we do not normally have our cherished memories so rudely shaken from us. Still, psychologists have shown in many studies over the past 40 years that our memories are surprisingly malleable. We can often remember things quite differently from the way they happened or, as in Piaget's case, have vivid, detailed memories of events that never happened at all. We next consider a laboratory paradigm that captures the effect in an easily studied manner.

## Associative memory illusions: the DRM effect

You have probably had the experience of listening to a story or a lecture and then something the speaker said led you to think further on the topic before your attention snapped back to the speaker. Later, if you are trying to recount the story to a friend, you might begin to relate a detail and then stop and wonder: did the speaker say that or was that something I thought while listening to her? Or, worse yet, maybe you never even wonder,
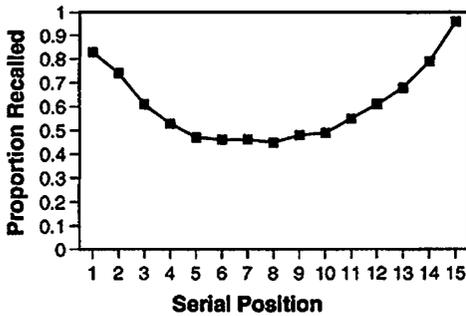
**Figure 8.2** Mean proportion recall of words as a function of list position (data are averaged over 24 lists). Data are from Roediger and McDermott (1995, Experiment 2).

but you confidently assert that she said something that you only thought. The events of life go whirling by, and the memorial residue is some combination of what really happened with how we recoded the information given our own background, interests and proclivities (Bartlett, 1932). The problem of discerning which of our thoughts came from real events and which were ones we inferred or imagined is called 'reality monitoring' (Johnson and Raye, 1981).

Roediger and McDermott (1995) developed a paradigm, first introduced by Deese (1959) for other purposes, to get at these issues. Subjects heard lists of words such as *hard, light, pillow, plush, loud, cotton, fur, touch, fluffy, feather, furry, downy, kitten, skin, tender.* They were instructed to listen carefully and, immediately after hearing the list, to write down all the words they could recall in any order they wanted (free recall). They were told to be very careful and to recall only items that they had just heard. The subjects' recall is shown in Figure 8.2 plotted against the input position of the words in the list (the data are averaged over 24 lists and many subjects). The figure reveals a standard U-shaped serial position function: subjects recalled the most words from the beginning of the series (the primacy effect) and from the end (the recency effect) – a standard finding. However, something unusual occurred in this experiment. When recalling each list, subjects tended to recall a particular word. The recalled word for this sample list was *soft*; in fact, the list was generated from the 15 words most closely associated from the word *soft* in norms of word association. That is, if students are given the word *soft* and asked to think of the first word that comes to mind, the 15 words in the list are the 15 most popular associates. Interestingly, subjects recalled the associated word that was never presented 55% of the time – about the same level or even slightly higher than recall of words that actually were presented in the middle of the list (like *cotton, fur*, etc. in this list). This illusory recall of a word not presented in a list (but strongly associated to the ones that were presented) is called the DRM (for Deese–Roediger–McDermott) effect.

Unlike Piaget's false memory, this one (albeit much more prosaic) appears immediately after study, with no appreciable delay between study of the list and its test. But is memory for the non-event rich and detailed? Do people really remember it? Roediger and McDermott (1995, Experiment 2) asked this question by giving a recognition test after the recall test. Subjects looked at words that were studied (*hard, cotton*) and words that were not studied in any list (*eagle, typhoon*) and, critically, the word implied by the list but not actually studied (*soft*). They asked the subjects to judge each word on the test as old (studied) or new (non-studied). If a word was judged old, subjects were asked to make a second judgement: did they really remember the moment the word occurred in the list
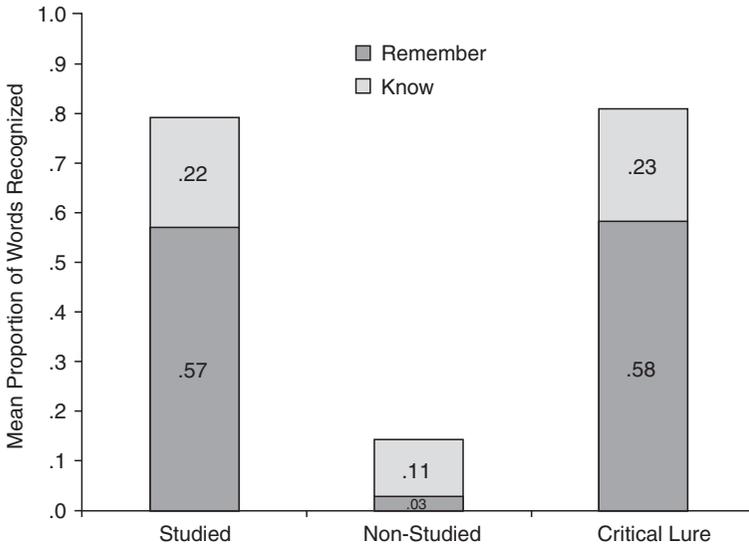
**Figure 8.3** Mean proportion recognition for studied, non-studied and critical lure words broken down as a function of remember/know judgements. Data are from Roediger and McDermott (1995, Experiment 2).

(e.g. the sound of a person's voice, the words before or after it), or did they just know it was on the list (but they could not remember the moment of presentation)? Tulving (1985) and others (Gardiner, 1988; Rajaram, 1993) developed this remember/know procedure to analyse a person's subjective experience during retrieval.

The results from the recognition test are shown in Figure 8.3. Studied words were recognized (called old) about 79% of the time; further, for those words called old, about 75% of the time subjects reported remembering the moment of the item's occurrence by providing a *remember* judgement. For the non-studied and unrelated words (the standard kind of lures on most recognition tests), the results were quite different. Subjects rarely called these items old (10% or so), and when they did make this mistake, they nearly always judged the item to be known and not remembered. After all, the item was not studied, so how could someone (just a few minutes later) have a strong experience of false remembering? The answer to this question lies in the bar on the far right in Figure 8.3: when the test item was strongly associated to one of the studied lists (like *soft* in our example list), a vivid false memory occurred. Subjects called such items old 81% of the time and, even more remarkably, they said they remembered the occurrence of the word in the list about 75% of those times. In fact, the results for the associated lure items like *soft* show about the same performance as items from the lists that were actually studied!

Of course, this laboratory sort of false memory does not rival the Piaget anecdote in its sweep and scale. Nonetheless, it provides a carefully controlled procedure by which genesis of false memories can be studied. One prominent theory to account for the effect (at a psychological level) is the activation-monitoring framework (Roediger *et al.*, 2001a; Balota *et al.*, 1999). Briefly, the idea is that the list of associates sparks thoughts (conscious or unconscious) of associated words (so people hear *hard, fur, cotton* and the word *soft* becomes highly activated). Once an item has been activated, then the subject has a reality-monitoring problem when retrieving words during the test, asking: 'Was this word presented or is it activated for some other reason?' Subjects often fail this reality monitoring test and report or recognize the associated word like *soft* as though it had actually been presented.

A large amount of research has grown up about this DRM paradigm, and Gallo (2006) has written an entire book towards understanding it and related phenomena. The DRM effect is large, persistent and robust across many conditions and subject groups. It provides a compelling and perplexing experience for all who try it. Piaget's false memory presumably developed over the years and we can understand how someone may not recall his childhood accurately so many years later, but the DRM illusion (which also involves remembering concrete details of an event that never happened) develops over seconds. The power of the demonstration, and our surprise in seeing that our recall is wrong, present a paradox to our understanding of how memory works. Of course, we are not arguing that the mechanisms of the Piaget false memory and those of the DRM illusion are the same; they surely are not. However, both phenomena indicate how people can remember events that never occurred.

## The curse of knowledge

The brilliant statistician, a leader in his field, is assigned to teach introductory statistics. This should be a breeze, he thinks. However, he is confounded by his class; they know nothing, they cannot understand anything he says. The students are similarly confounded. The teacher talks in equations, does not give concrete examples, and seems to assume that they have already had several statistics courses. They have not. The professor in this instance is hampered by the curse of knowledge – he knows so much about the field that he can no longer put himself in the place of a student in college who has never had a statistics course.

Similarly, computers and all sorts of other technologies (think of your TV, DVD player and cable box) are designed by electrical engineers, computer scientists and others of their ilk. The early personal computers were maddeningly difficult to use, and sales suffered. The reason was that engineers designed them so other engineers could use them – not normal people with no engineering background. Apple and some other far-sighted computer companies started hiring human factors psychologists to help engineers to redesign the computer to take people – the human factor – into account. The psychologists had to get the engineers to overcome the curse of their knowledge and make the equipment so that nearly any slob could use it.

Elizabeth Newton (1990) conducted an experiment that reveals the curse of knowledge. She made up a list of tunes that nearly every American grows up knowing – Happy Birthday to You; Shave and a Haircut, Two Bits; The Star Spangled Banner, etc. Two students sat on opposite sides of a screen in a room unable to see each other, though they could hear each other, with the screen on a desk between them. One student, the sender, tapped out a given tune with their knuckles on the table. The sender had to judge whether or not his/her performance was successful in revealing the song to the other person, while the receiver's job was to guess the identity of the song from a list of 20.

The senders, as a group, seemed relatively modest. They thought that the receivers would be able to identify the tune they played about 50% of the time. However, the receivers were able to identify the tune correctly only 3% of the time, just at chance levels! The senders were actually wildly optimistic. One plausible reason is the curse of knowledge: when the sender was tapping her knuckles on the table, she was mentally hearing the music and words of 'Happy Birthday' or some other song. This vivid imagery made her sure she was tapping out a great song, but of course what the receiver was hearing was some knocks on the table (not the music, not the words). The sender could not appreciate how difficult a
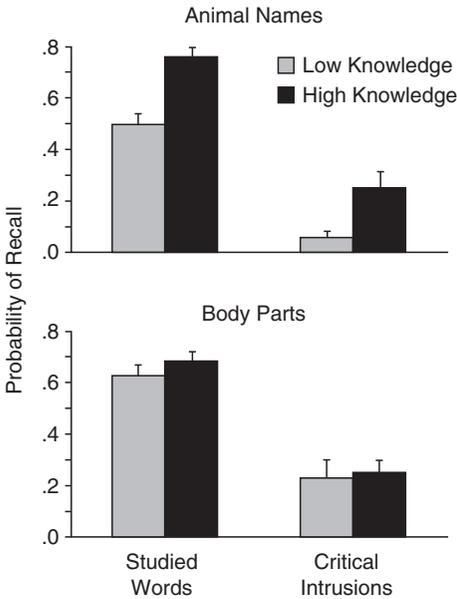
**Figure 8.4** Recall of animal names (top panel) and body parts (bottom panel) as a function of football knowledge (high versus low). Data are from Castel *et al.* (2007).

job the receiver faced because she was cursed with knowledge of the song. Similarly, the statistics professor giving his lecture can imagine that he is making brilliant connections among topics, dazzling the students with his knowledge and erudition. However, the students do not know enough to be dazzled; they are hearing a lot of jargon that bounces off them rather than being absorbed. They do not have the knowledge structures (the schema, to use jargon from cognitive psychology) that would permit the lecture to be understood.

The curse of knowledge can also show up in DRM-type memory studies. Castel *et al.* (2007) tested students who were either avid and knowledgeable football fans or who were not. They gave them two lists of material to remember. One list was animal names, but the names all belonged to professional US football teams (*dolphins, broncos, falcons, colts, jaguars, bengals, seahawks, rams, lions, ravens* and *bears*), but some other team names were omitted (*eagles, panthers* and *cardinals*). The other list was composed of body parts, and again 11 items were presented (*arm, knee, mouth, stomach*, etc.) and three common items were omitted from the list (*leg, head, nose*). Thus, the lists are like DRM lists in that they both cluster about a theme (animals, body parts) with some items presented and some omitted. Students studied both lists and, after a 10-minute delay, tried to recall each one when cued with the category name (animals or body parts).

The results are shown in Figure 8.4, with recall of animal names shown at the top and body parts shown below. The bars represent the students who were either high or low in football knowledge. Those with great football knowledge recalled the animal names (belonging to team mascots) better than those students who did not know as much about football. This pattern shows the positive effect of expertise, of a case where knowledge is not a curse but a blessing. However, notice that there is a downside, too. Those with high football knowledge were also more likely to falsely recall animal mascot names that were not on the list (see the bars labeled critical intrusions). Thus, increased knowledge cursed these students with higher levels of false memories.

However, maybe those students with high football knowledge were just different somehow from the other students. That possibility seems unlikely as judged by the other data in Figure 8.4. When recalling body parts, both groups looked quite similar in terms of both accurate and false recall, so there was probably no general difference in ability across groups. The difference between the groups was in terms of expertise for football – this knowledge was both a blessing, in helping recall of animal names actually presented, and a curse in promoting false recall (see Chapter 9 for discussion on the curse of expert knowledge).

## The knew-it-all-along effect (hindsight bias)

This common illusion is a cousin of the curse of knowledge, or perhaps a species of it. People have great confidence (after the fact) that they knew something (or could have predicted something) when in fact they could not have. This bias shows the value of the proverb that 'hindsight is 20/20'. Foresight is typically myopic. As we write this chapter, many books are appearing claiming that the factors that caused the economy to crash in 2007–2008 were huge (mortgage risks, all kinds of risky investments built on unsound mortgages and so on; e.g. Foster and Magdoff, 2009). However, all these factors were clear before the crash and yet practically no one predicted it. In hindsight, the crash and its causes seem obvious, but no one in power displayed the foresight to identify and prevent them.

A laboratory paradigm to identify and study hindsight bias was identified in two important papers by Baruch Fischhoff in 1975 (Fischhoff, 1975; Fischhoff and Beyth, 1975). Students were told they were to assess the likelihood of the outcome of events. An event was described, they were provided with four possible outcomes and they had to assign probabilities for each possible outcome. In one study, students read brief passages (about 150 words) describing a historical or clinical event that was true but would be unknown to most of them. For example, one incident was about a battle in 1814 in India between the British and Gurkas of Nepal. After the description was given, students were given the four outcomes (the British won; the Gurkas won; a military stalemate ensued; or a military stalemate occurred followed by a peace treaty). Two groups of students were given the same description, but with one difference. One group was told how the event actually came out in the last sentence of the paragraph, while the other was not. The students' task was to assign probabilities to the four possible outcomes so that they would sum to 100. After that, they justified their responses by saying which parts of the passage were most relevant in making their judgements.

The basic finding, which has been replicated many times, is that students who knew the outcome deemed it much more probable than students who were not told the outcome. When students know an outcome, they selectively choose the evidence to justify why they thought they would have predicted it. However, students in the other group who did not know the outcome and actually did have to predict it generally did not arrive at the same conclusion. Their probabilities were more evenly split among the alternatives. In other studies, Fischhoff (1975) demonstrated that people are generally unaware of this hindsight bias. They fully believe that they could have predicted the events given the other information in the paragraph, even though the control groups show that this is not so. In a later replication and extension of this work, Wood (1978) provided another descriptive label – the 'knew-it-all-along effect'. Keep this in mind as you hear media figures or friends

pontificate on why the stock market did what it did during the day, why the President acted as he did, and so on. In the case of the stock market, you can ask: if you saw this trend coming so clearly, why didn't you get rich?

## Paradoxes of difficulty

People want learning to be quick and easy. If you need to be convinced that this statement is true, then do a quick Google search for 'learning' paired with 'fast' (or any synonym of your choosing). Among the millions of website hits that result, you will find products, programs and other tools that all claim to speed up the learning process. Why spend a year living abroad in Peru when you can learn to speak fluent Spanish in mere weeks through a language-learning program like Rosetta Stone? Why read *A Midsummer Night's Dream* when you can learn all about Shakespeare's romantic comedy by consulting a study guide like Cliff's Notes? Obviously, these hypothetical questions ignore the richness of learning that accompanies the experience of living in a foreign country or reading a literary classic. Yet, they raise an important point: if the goal is to attain some criterion level of learning (e.g. fluency in speaking Spanish, knowledge of the plot of *A Midsummer Night's Dream*, etc.), why not do so in the quickest and easiest way possible?

The problem with this line of reasoning is that the level of performance during learning is a poor indicator of whether the knowledge or skill will be retained over longer periods of time. That is, reaching some criterion level of performance during learning does not guarantee that the knowledge or skill will be well remembered in the future. As Bjork and Bjork (1992) have argued, performance during learning reflects momentary accessibility of knowledge or skill (i.e. *retrieval strength*), rather than how well it has been stored in memory (i.e. *storage strength*). When retrieval strength is high, but storage strength is low, performance will be excellent in the short term, but it will suffer in the long term. An example would be remembering what you had for breakfast this morning (assuming you don't have the same breakfast every morning) – you will have no trouble retrieving that information today, but you would probably fail to retrieve it if you tried again a month from now. Thus, the key to ensuring that knowledge or skill is retained over the long term is building up high levels of storage strength.

Based on this analysis, R.A. Bjork and colleagues (Bjork, 1994a, 1994b; Christina and Bjork, 1991; Schmidt and Bjork, 1992) proposed the paradoxical concept of 'desirable difficulties' in learning. They argued that introducing difficulties during learning can actually increase long-term retention because the greatest gains in storage strength occur when retrieval strength is low. In other words, successfully retrieving information under difficult circumstances will lead to greater increments of storage strength relative to retrieving that information under easy circumstances. For example, imagine you are introduced to someone new at a cocktail party – if you retrieve that person's name immediately (i.e. when retrieval strength is high), the gains in storage strength will be much smaller than if you retrieve the name after 5 minutes of conversation (i.e. when retrieval strength is low). The idea of 'desirable difficulties' in learning is paradoxical in that it contradicts the commonly held belief that factors which enhance performance or speed improvement during learning also produce superior long-term retention. We now turn to three examples of such 'desirable difficulties' in learning, each of which could be considered paradoxical as well.
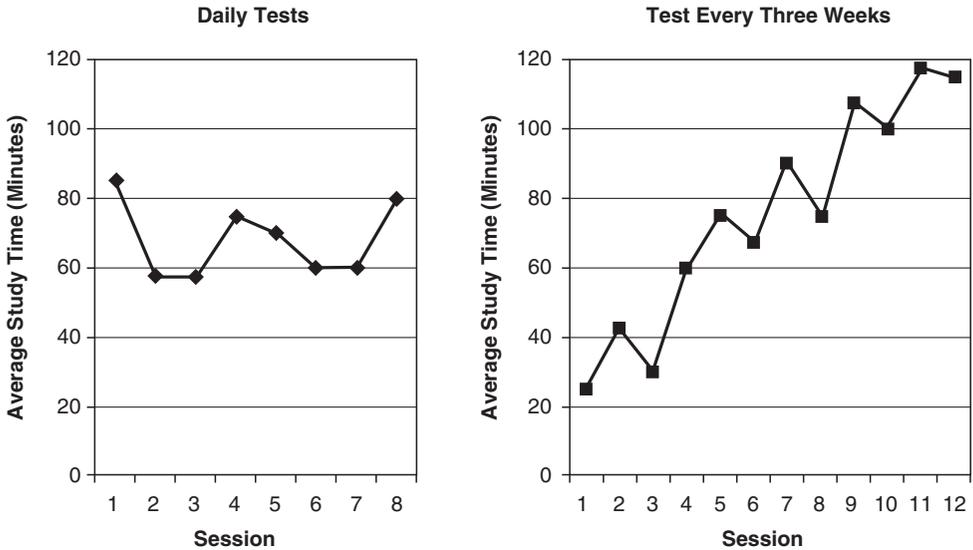
**Figure 8.5** Average number of minutes spent studying per session in the daily test (left panel) and 3-week test (right panel) conditions. Data are from Mawhinney *et al.* (1971).

## Spacing

Mark Twain once said, 'never put off till tomorrow what you can do the day after tomorrow'. People are masters at the art of procrastination, especially when it comes studying for a test. As anyone reading this book can readily attest, procrastination inevitably leads to cramming. In general, students spend relatively little time studying until immedi-ately before the test, a pattern of behaviour that has been referred to as the 'procrastination scallop' (Michael, 1991). Figure 8.5 shows the results of an experiment conducted by Mawhinney *et al.* (1971) in which students were either tested every day or every three weeks. In the daily testing condition (left panel), the average number of minutes that students spent studying each day remained high and relatively constant; however, when tests were given every three weeks (right panel), the students' study behaviour exhibited the usual scalloping pattern leading up to the test that occurred after the twelfth study session.

Of course, cramming is a perfectly good way to maximize performance on an immediate test, but much of the information is quickly forgotten after the test. If long-term retention is the goal (which is certainly true of formal education generally, but perhaps not for students individually), then it is much better to space out or distribute study over time. The mnemonic benefit of spaced practice over massed practice (i.e. cramming) is one of the most robust and well-replicated findings in research on human memory and learning (Glenberg, 1976; Melton, 1970; for a review, see Cepeda *et al.*, 2006; Dempster, 1989). Indeed, the spacing effect, as it is often called, is also one of the oldest findings – it was described by Ebbinghaus ([1885]1967) in the first experimental investi-gation of human memory.

Spaced practice constitutes a 'desirable difficulty' in that it takes longer to reach the criterion level of performance during learning relative to massed practice, but it leads to better long-term retention. Prior research has found spacing effects for inter-study intervals ranging from several seconds (see Underwood, 1961) to years (e.g. Bahrick

*et al.*, 1993). However, a recent meta-analysis performed by Cepeda and colleagues (2006) showed that the optimal spacing interval seems to depend on how long the information needs to be retained: final test performance is maximized when the inter-study interval is roughly 20% of the retention interval. Although subsequent research suggests that the retention-maximizing ratio may vary slightly depending on the retention interval (see Cepeda *et al.*, 2008), a general rule of thumb is that greater spacing will be beneficial for longer retention intervals.

## Variability

When people are learning to perform a task or attempting to acquire some knowledge, often they will repeatedly practice the same action or study the same material. Children who are learning to write in cursive often practise writing one letter many times before moving on to the next letter. Professional basketball players spend a great deal of time practising how to shoot a free throw. During an exam period, students often concentrate on studying for one exam before moving on to study for another exam. Consistent practice in massed fashion often enables learners to reach some criterion of performance very quickly. However, as we discuss above, fast learning does not always lead to good long-term retention.

Although consistently practising the same action or studying the same material may have benefits in certain circumstances, many studies show that introducing variability during practice can produce superior performance on later tests. For example, Kerr and Booth (1978) conducted an experiment in which two groups of 8-year-old children practised throwing beanbags at targets during a learning phase. In the variable group, the children threw beanbags at targets that were 2 feet and 4 feet away. In the constant group, they threw beanbags at targets that were 3 feet away. After completing the learning phase, the children received a final test on a target placed 3 feet away. It would be easy to assume that the constant group would perform better on this final test than the variable group because the constant group had practiced this distance, but the variable group had not. However, contrary to one's intuition, it was the variable group that produced the superior final test performance. Indeed, many other studies of motor skill learning have shown a benefit of variable practice relative to a consistent practice schedule (for review see Shapiro and Schmidt, 1982).

Interestingly, the benefits of introducing variability during learning that occur in motor learning seem to hold for the learning of verbal information as well. Goode *et al.* (2008) compared variable and consistent practice in an experiment that involved solving anagrams. In the 'same' condition, students had to repeatedly solve a set of anagrams, each of which was always presented in the same way (e.g. LDOOF; to which the answer is FLOOD). In the 'varied' condition, they also repeatedly solved a set of anagrams, but they received a different variation of the anagram (e.g. FOLOD, OOFLD, and DOLOF) each time they had to solve it. Once the practice phase was completed, students were tested with the variation of the anagram that was practised three times in the same condition, but never practised in the varied condition. As Figure 8.6 shows, students who received varied practice produced a greater proportion of correct solutions to the anagrams on the final test relative to those who had practised with the same variation of anagram.

Introducing variability during learning can be considered a 'desirable difficulty' in that variability often slows learning and requires greater effort on the part of the learner; however, as the findings described above demonstrate, variable practice often produces
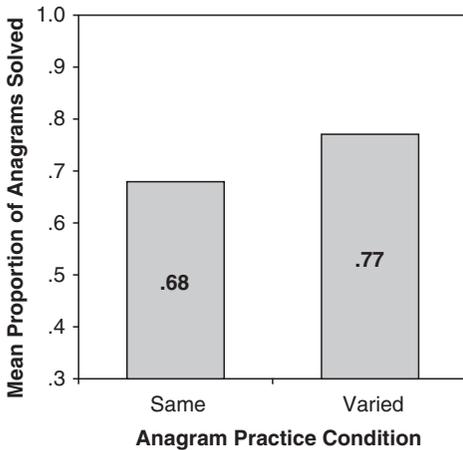
**Figure 8.6** Mean proportion of correct solutions to the anagrams on the final test in the same and varied practice conditions. Data are from Goode *et al.* (2008).

better subsequent performance than consistent practice. A related idea is the concept of interleaving tasks during learning. If one must learn to perform three tasks (A, B and C) during a fixed period of time (e.g. a 3-hour training session), then one must decide how to distribute practice on these three tasks. One possibility is to practise Task A for an hour, then practise Task B for an hour, and so on for Task C. However, one could also practise each task for 10 minutes at a time, continuing to switch between tasks over the 3 hours. Research has shown that the latter schedule of practice, often called interleaving, leads to better subsequent performance (e.g. Shea and Morgan, 1979; for review see Magill and Hall, 1990). The benefits of interleaving are likely derived in part from the effects of spacing practice (as discussed above), but presumably these benefits also result from introducing variability during practice.

## Testing

What activities produce learning? The first answer that comes to mind for most people is probably studying. Indeed, when students are asked about their study habits, they often report that their top strategy for learning is to repeatedly read information (e.g. Karpicke *et al.*, 2009; see too Kornell and Bjork, 2007). In contrast, testing is an answer that would be at the end of most people's list, if they include it at all. One reason that testing is likely to be omitted from a list of activities that produce learning is that people generally conceptualize testing as an assessment tool. That is, testing is assumed to be a neutral event in which knowledge is assessed without changing memory, much as stepping on a scale does not alter a person's weight. However, research on memory and learning has shown that the act of retrieving information from memory actually changes memory (e.g. Bjork, 1975), often leading to better retention over time (e.g. Carrier and Pashler, 1992).

The finding that practice in retrieving information from memory (i.e. testing) produces superior long-term retention is commonly referred to as *the testing effect* (for review see Roediger and Karpicke, 2006a). Critically, the mnemonic benefits of retrieval practice emerge even when neither feedback nor further study opportunities are provided and when compared to a control condition that re-studies the information for an equivalent amount of time (e.g. Glover, 1989; Roediger and Karpicke, 2006b). An experiment by Karpicke and Roediger (2008) provides a simple, yet powerful illustration of this robust phenomenon. In
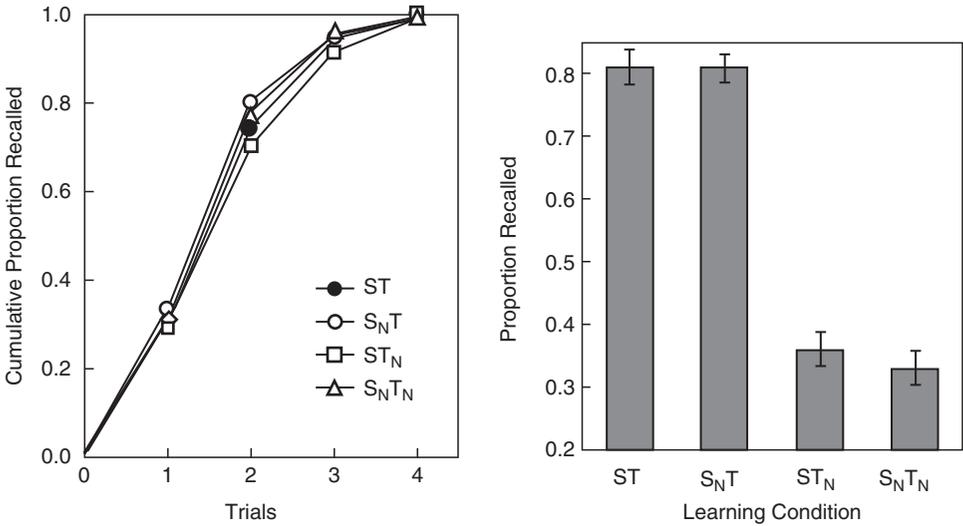
**Figure 8.7** Proportion of correct responses on the initial learning tests (left panel) and the delayed final test (right panel). Data are from Karpicke and Roediger (2008).

a first phase, all students studied a list of 40 Swahili–English word pairs (e.g. mashua–boat) and then they were tested on each pair from the list (e.g. mashua– ?). When a student successfully recalled the English word on a test trial, then that pair moved into a second phase that consisted of four additional trials in one of four experimental conditions. In the 'standard' (ST) condition, students repeatedly studied and took tests on all the word pairs. In the 'repeated testing' condition, they took tests on the word pairs, but they did not study them any more ($S_NT$). In the 'repeated study' condition, they studied the pairs, but did not take any more tests ($ST_N$). Finally, in the 'drop' condition, students neither studied nor took a test again ($S_NT_N$).

The left panel of Figure 8.7 shows the cumulative learning curves (giving credit for recall the first time an item was recalled and ignoring repeated recall of the same item in some conditions) for all four experimental conditions. The rate of learning did not differ among the conditions, and every student had successfully recalled the correct English word for each pair in the list by the end of the learning session. Thus, if performance during initial learning is the criterion by which the efficacy of any learning strategy is judged, then all four experimental conditions would be assumed to be equally effective. However, this assessment changes drastically when long-term retention of the word pairs is considered. As the right panel of Figure 8.7 shows, the pattern of performance on a final test given one week later was very different. Students in the repeated study and drop conditions recalled 36 and 33%, respectively – a substantial decline when compared with performance at the end of learning. In contrast, students in the standard and repeated testing conditions recalled approximately 80% of the word pairs, which means there was relatively little forgetting in these conditions. What was the critical difference between these conditions? Retrieval practice. In the standard and repeated testing conditions, students continued to be tested, but they did not take any further tests in the repeated study and drop conditions.

Returning to the central theme of this section, testing clearly represents a 'desirable difficulty' in learning. The act of retrieving information from memory requires greater

effort than passively studying that information, but this difficulty during learning leads to superior performance over the long term. The mnemonic benefits of retrieval practice are particularly interesting in light of the fact that most people do not think of testing as a way to promote learning. Rather, they report studying as their top learning strategy. Looking at the results of the Karpicke and Roediger (2008) experiment, it is important to note that the standard condition produced equivalent retention to the repeated testing condition, even though the standard condition contained many more study trials. This result suggests that continuing to study an item once it has been recalled does little to improve retention. It also fits nicely with the findings of other studies which have shown that after an initial reading of the material, re-reading it produces relatively limited memorial benefits (see Callender and McDaniel, 2009).

## Paradoxes of interference

Interference from other events is perhaps the most potent cause of forgetting any particular event. Suppose you are a frequent traveller and you are asked to remember in great detail the airplane trip you took five flights ago. The four flights since that trip would provide retroactive interference as you tried to retrieve the critical trip, and all the trips you took before that critical trip (but especially the ones immediately before it) would provide proactive interference. The difficulty is to hone in on the particular event and ignore ones like it that occurred before it (creating proactive interference) and after it (creating retroactive interference). Interference occurs in many forms, but we will concentrate on three interesting interference phenomena in this section of the chapter.

## The misinformation effect

Loftus and Palmer (1974) were interested in interference in a situation that has great implications. When a person is a witness to a crime, they may have to testify in a court of law months (or even years) later about what they saw or heard. Can information that occurs after an event be incorporated into memory of the event? Can the interfering information change or override what one actually saw? Loftus and Palmer (1974), and many other researchers since then, have shown that the answer to this question is yes.

The typical misinformation experiment involves three stages. First, a person witnesses a simulated crime (e.g. a repairman fixing a desk in an office steals money from a wallet). During a second phase, the person reads a report ostensibly produced by some other witness to the event, but there are errors in the report. For example, if the thief had been using a screwdriver to repair the desk, the report might refer to him using pliers. The third part of the experiment involves a test in which questions are asked about the original event (sometimes with subjects being warned that the reports they read might have errors). The subject might be asked 'What tool did the man use to fix the desk?' The finding is that the misinformation presented in the post-event report can alter the subjects' response. Relative to a control condition in which no misinformation is presented, subjects will recall items suggested in the report as actually having been present in the scene (Loftus *et al.*, 1978). Further, people will often say they actually remember the item in the scene (Roediger *et al.*, 1996), using Tulving's (1985) remember/know procedure.

The fact that information which is presented after an event can alter the memory of that event has enormous implications for the legal profession and the veracity of eyewitness testimony. After all, once a person has witnessed a crime, he/she will think about and
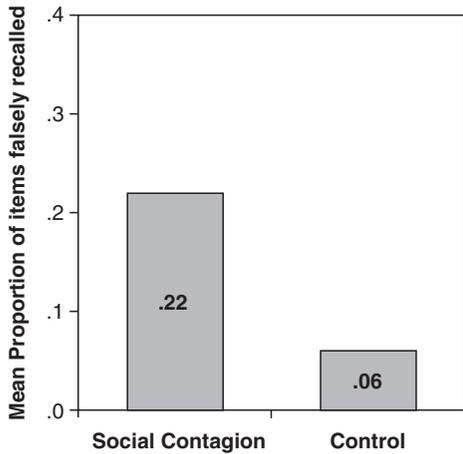
**Figure 8.8** Proportion of items falsely recalled for the social contagion and control conditions. Data are from Roediger *et al.* (2001b).

recount the event in response to questions from police, friends and lawyers long before testifying in court. Each remembrance of the event may (depending on the context, the questions asked, etc.) alter it in subtle ways.

The misinformation provided in the experiments just discussed was from a written source. Roediger *et al.* (2001b) developed a paradigm to see whether similar effects would occur when students took turns recalling items from a scene. The two people came to the lab together and watched the same scenes. However, one person was an experimental confederate. The pair watched 6 scenes together and then took turns recalling 12 items from the scenes, 6 apiece. The confederate accurately recalled 6 items from 3 scenes, but for the other 3 scenes he recalled 4 correctly and got 2 wrong (producing items that might have been in the scene but were not). Finally, at the end of the experiment, the students were separated and tested separately (of course, only the real subject got the test). The subject was told to try to recall as accurately as possible the items that were in the original 6 scenes as the scenes were cued one by one. The interest was in whether the subjects would 'remember' items as being in the scene that the confederate suggested (and which were not in the scene). The answer is Yes, and the results are shown in Figure 8.8. The authors called this effect 'the social contagion of memory', because the confederate's erroneous memories 'infected' those of the subject. This effect is another example of how retroactive interference can create illusory memories.

In both the standard misinformation effect experiment and the social contagion variation on the theme, people are unaware of how their memories have been affected by information occurring after the events. As noted above, they often display high confidence in their illusory memories and claim to remember them in Tulving's (1985) special sense of the term. Warnings that the report includes errors or that the confederate made mistakes weaken but do not eliminate the effect (e.g. Meade and Roediger, 2002). Thus, as in the DRM effect, we are left in the paradoxical state where people cannot distinguish real events they actually experienced or saw from ones that were merely suggested, which leads to yet another paradox of memory. Of course, the misremembered events in the cases above were ones that were suggested to have been in the scene. Surely people would not misremember things they actually did (not just things they saw). Or would they?

## Imagination inflation

Forming images is an age-old way to improve memory for verbal items (e.g. Bower and Reitman, 1972). Many mnemonic devices depend on imagery (e.g. Roediger, 1980). However, imagining events can also create errors. In experiments by Johnson and colleagues (1979), subjects either saw pictures of a butterfly various numbers of times or saw the word butterfly various numbers of times. Later they were asked to judge the frequency that a particular word or picture was observed. Subjects were generally fairly accurate at this task. However, in a condition in which subjects saw words but were asked to form mental images of the words' referents (see the word butterfly but form a mental image of a butterfly), the authors showed that the estimates of having seen the pictures of butterflies were increased. That is, subjects confused some of their own images for actually occurring pictures.

Goff and Roediger (1998) asked whether a similar effect would occur with behaviours people performed. Based on earlier work on action memory involving 'subject-performed tasks', they developed a set of 60 tasks that people could do while sitting at a desk either with small objects ('pick up the paper clip') or with hand movements ('touch your left ear with your right hand'). On a first day in the experiment, students heard instructions to perform tasks like these between one and six times; they performed some tasks but only listened to the instructions for others. They came back on a second day for a second session in which they now imagined performing tasks. Some of the tasks were ones they performed the first day, whereas others were new. Again, they were asked to imagine performing the tasks from one to six times. Then the students had a long break. They came back to the lab two weeks later. They were told that they would be given a test for the events they actually performed on the first day. They were told to ignore any events they heard on the second day and to concentrate on remembering only what happened on the first day.

The results showed that, despite the instructions, the act of imagining events on Day 2 inflated the judgements of the number of times they were performed on Day 1. This was true for events that actually had been performed on Day 1, but more importantly, it was true for events that were not performed on Day 1. That is, repeatedly imagining doing something made people believe they had actually done it. This effect has been called 'imagination inflation' and is studied in the laboratory, as here, or in more natural settings with events from childhood. In this latter case, imagining events from childhood (e.g. running through the house, slipping, and cutting one's hand on broken glass in a window) increased the probability that people thought the events had actually happened to them. Thomas *et al.* (2003) showed that instructing people to vividly imagine events (colours, sounds, etc.) increased imagination inflation. Imagination inflation shows that when we imagine ourselves performing actions we may later believe we actually did the action. Once again, our imaginations can play tricks on our memories.

## The self-limiting nature of retrieval

Think back to a recent time that you were together with a large group of people – maybe 10 to 15 individuals – all of whom you know. It could be a dinner party, a work meeting, or any other type of gathering. Now try to list the names of all the people at this event. The first couple of names will be easy; however, you will probably find that it becomes progressively harder as you continue to retrieve names from memory. You may even fail
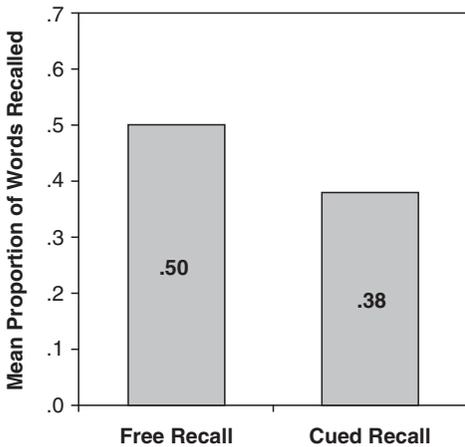
**Figure 8.9** Mean proportion recall for the free recall and cued recall conditions. Data are from Slamecka (1968, Experiment 2).

to remember the last two or three people – a frustration that we have all experienced before. This phenomenon presents a paradox because the retrieval of some information is generally thought to facilitate the retrieval of associated information. For example, the use of mnemonic devices, such as the peg-word method, is predicated on the idea that associations between memories enhance retrieval (Roediger, 1980). Why, then, does remembering sometimes grow harder as we successfully remember more related information?

One potential answer to this question is that, under certain circumstances, recall is a self-limiting process because the act of retrieving a memory can interfere with the retrieval of related memories, a phenomenon that is often referred to as 'output interference' (see Roediger, 1974, 1978; Tulving and Arbuckle, 1963) and has more recently been called retrieval-induced forgetting (Anderson *et al.*, 1994). The part-list (or part-set) cueing paradigm is one method that has been used in the laboratory to investigate output interference. In a typical experiment, subjects study a list of words and then they attempt to recall all the words in the list or they are 'cued' with a subset of the words in the list and must recall the other words in the list. For example, Slamecka (1968, Experiment 2) showed subjects a list of 30 words and then had them either recall all 30 words (free recall) or presented them with 15 of the words and had them recall the other 15 words (cued recall). Figure 8.9 shows the mean proportion recall in each group for the 15 words that were not presented in the cued recall condition. As you can see, subjects in the free recall group recalled a significantly greater proportion of the words relative to the cued recall group. That is, the presence of some of the words from the list seems to interfere with recall of the remaining words.

Why is recall a self-limiting process? Rundus (1973) offered an explanation in which knowledge is conceptualized as a hierarchical structure with groups of associated items connected to a common node (e.g. a category label). When recalling items associated with a given node, those items are retrieved through a process of sampling with replacement. The act of retrieving an item strengthens that item, which increases the probability that that item will be recalled again in the future. Thus, as more items from the set are recalled, these 'old' items begin to be repeatedly recalled, preventing the recovery of non-recalled items from the same set. Several other theories of these effects also exist (see Bäuml, 2008).

## Paradoxes caused by fluency of cognitive processing

Try to remember a long trip that you took as a child. Perhaps you went to explore another country, or to visit relatives that live far away, or maybe you travelled to attend some unique event. Now, ask yourself, 'What makes me sure that I really experienced this event and that it is not a product of my imagination?' One way of answering this question is to evaluate the contents of the memory. Real memories tend to contain more idiosyncratic details, a greater amount of sensory, spatial or temporal information, and vivid imagery relative to imagined events (see Johnson *et al.*, 1993). In addition to the objective contents of the memory, people often rely upon the subjective experience of remembering in order to make such a determination. The act of remembering a past experience is often accompanied by a feeling of familiarity, which is interpreted as a signal that the memory truly represents a past experience. In the absence of this subjective experience, we may know that a particular experience occurred, but the feeling of ownership is lost.

One idea is that the subjective experience of remembering is derived directly from the memory trace. However, a problem with this idea is that the act of retrieving a memory does not always give rise to a feeling of remembering. For example, people with amnesia utilize representations of past experiences to facilitate performance on implicit memory tasks, but do not 'remember' those past experiences (see Roediger, 1990). In addition, people can experience a feeling of remembering in the absence of a memory trace – such as when an amnesic patient confabulates by making up a false response to a question but strongly believes it to be true. Thus, the existence of a memory trace is neither necessary nor sufficient for a person to experience a feeling of remembering.

An alternative idea, proposed by Jacoby and colleagues (1989), is that the subjective experience of remembering results from an attribution or inference about the fluency of cognitive processing. For example, when we re-read an old book, we process the prose more fluently and we (correctly) attribute that ease of processing to having read the book before. Similarly, when you tried to remember a long trip during childhood a few moments ago, the full memory likely came to mind relatively quickly after you had identified it and you (correctly) inferred that this fluency was due to you having previously experienced this event. Of course, these two examples illustrate instances in which fluent cognitive processing results from the existence of a memory trace and that fluency is correctly attributed to the memory trace. However, many other factors can also influence the fluency of cognitive processing – the use of overly complex words in prose can decrease perceptual fluency (e.g. Oppenheimer, 2006), priming people with the answer before displaying a question can increase the retrieval fluency (Kelley and Lindsay, 1993).

In addition, the attributions that people make about the origin of such fluent processing are often driven by the goals of the ongoing task. Increased fluency may be attributed (either correctly or incorrectly) to a prior experience if the current goal is to remember, but it may be attributed (again either correctly or incorrectly) to another factor if remembering is not the current goal. For example, if you are evaluating a piece of writing that is printed in a hard-to-read font, you might (incorrectly) attribute the decrease in perceptual fluency to the quality of writing and give it a poor evaluation (see Oppenheimer, 2006, Experiment 4). Thus, the accuracy of people's attributions about their subjective experience of remembering depends on both the source of the fluency of cognitive processing (prior experience versus other factors) and the current goal of the ongoing task (remembering versus another goal). Generally speaking, fluency is a good indicator of previous experience (i.e. the

retrieval of a memory trace often produces a feeling of familiarity), and we often interpret this fluency correctly. However, under certain circumstances, our reliance on fluency can paradoxically result in misattributions. We now turn to describing three examples of paradoxes caused by the fluency of cognitive processing.

## Cryptomnesia

Helen Keller was an American author and political activist, whose accomplishments are amazing because she was born deaf and blind. In 1892, an 11-year-old Keller published a story called *The Frost King*. Readers immediately noticed a striking similarity to a story called *The Frost Fairies* that appeared in a book written by Margaret Canby and published in 1874. Keller was accused of plagiarism – a charge that she vehemently denied. She claimed to have no recollection of being told *The Frost Fairies* story, but it later emerged that a family friend had communicated the story to her via her teacher, Anne Sullivan, tracing letters on her hand several years earlier. Made to stand trial before a tribunal of the Perkins Institute for the Blind, Keller was acquitted of intentional plagiarism in a close vote. The members of the tribunal who voted 'not guilty' were convinced that it was a case of cryptomnesia or unconscious plagiarism.

Cryptomnesia occurs when people retrieve other people's ideas and mistakenly believe that they generated them, either at that moment or at an earlier time. Helen Keller may have inadvertently plagiarized *The Frost Fairies* because it came to mind without any feeling of familiarity, and the absence of a subjective experience of remembering led her to believe that it was her own idea. Alternatively, if she did experience fluency in retrieving the story, she may have misattributed that fluency to the quality of the story because the goal of the ongoing task was creating a story (i.e. not remembering the previous experience of being told a story).

In the laboratory, cryptomnesia has been investigated using a paradigm in which two or more students collaborate on a generation task and then later try to remember who generated each idea and/or generate new ideas. For example, students might be asked to generate exemplars from categories (e.g. Brown and Murphy, 1989) or identify words in a word-search puzzle (e.g. Marsh and Bower, 1993). After the initial generation phase, they might have to recall the items that they generated earlier, generate new items, and/or take a recognition test that includes items that were generated by themselves and their partner as well as new items. In such experiments, cryptomnesia can occur in two ways: students can recall another person's item as their own, or they can generate a (seemingly) new item that was actually generated in the initial task.

Studies using this paradigm have shown that both types of unconscious plagiarism are quite common (e.g. Brown and Murphy, 1989; Marsh and Bower, 1993). The incidence of cryptomnesia increases when the final test phase is delayed rather than given immediately (e.g. Brown and Halliday, 1991). The goals of the ongoing task can also influence the amount of cryptomnesia observed. For example, a greater incidence of cryptomnesia is generally observed in tasks that involve generating new items relative to a recognition test in which students must categorize items as their own, someone else's or new (e.g. Marsh *et al.*, 1997). In addition, the incidence of cryptomnesia increases when people experience high incidental effort while working to generate items, but low effort when the solutions appear (Preston and Wegner, 2007). Presumably, this effect occurs because people misattribute the feeling of greater effort and subsequent release from effort to their own successful

generation of the item. Interestingly, recent research suggests that separate processes might give rise to the two types of plagiarism described above because certain manipulations (e.g. feedback on the quality of the ideas generated; Perfect and Stark, 2008) and individual difference variables (e.g. age; McCabe *et al.*, 2007) affect each type of plagiarism differently.

## Availability

Fluency can also affect people's judgements when relevant information comes to mind with ease. Consider the following question: Are you more likely to die from a car accident or a medical error? If you answered 'car accident', like many other people would, then you are wrong – assuming that you live in the United States of America, but it is probably the wrong answer in other countries too. While the number of fatalities from car accidents in USA has averaged between 40,000 and 45,000 each year over the past decade, medical error accounts for up to 225,000 deaths per year by one estimate (Starfield, 2000).

People make errors like this one because they often base their judgements on how easily relevant instances come to mind, a strategy that Tversky and Kahneman (1973, 1974) have called the 'availability heuristic'. When searching for relevant information on which to produce an answer to the question above, you likely had an easier time retrieving instances in which people died from a car crash than from medical error because the former is more prominently featured in the news, TV shows, books, etc. Reliance on the availability heuristic does not always result in an erroneous judgement because, generally speaking, availability is correlated with ecological frequency; however, it consistently leads to systematic biases under some circumstances.

In studies that investigate the circumstances under which people rely on the availability heuristic, the critical manipulation often involves a factor that affects the fluency or ease with which information comes to mind. For example, Carroll (1978) explored whether having people imagine the outcome of a future event would increase the availability of that outcome and thus bias subsequent judgements. In one experiment, he had people imagine either Jimmy Carter or Gerald Ford winning the upcoming US presidential election. When people were later asked to predict who was more likely to win the election, they tended to pick the candidate that they had imagined earlier. In another experiment, he had people imagine a good season or a bad season for the US college football team that had won the national championship during the prior year. When asked about whether that team would get a bowl bid (i.e. an invitation to participate in prestigious post-season game) at the end of the upcoming season, people were more likely to predict a bowl bid if they had imagined a good season.

Although factors that increase the availability of relevant information can influence judgements, the way in which such information is used depends upon the fluency or ease with which it is retrieved. Schwarz *et al.* (1991, Experiment 1) manipulated the number of examples that people had to generate about one of two types of behaviour – acting assertively or unassertively. Subjects recalled either 6 or 12 examples of situations in which they behaved assertively or unassertively. In the two groups that recalled 12 examples of assertive or unassertive behaviours, retrieval was difficult because it was hard to generate such a large number of examples. In contrast, retrieval was relatively easy for the two groups that recalled six examples of assertive and unassertive behaviours. Later, subjects were asked to answer some general questions as part of an unrelated task, including a question on which they had to evaluate their assertiveness on a 10-point scale (where 1 equalled
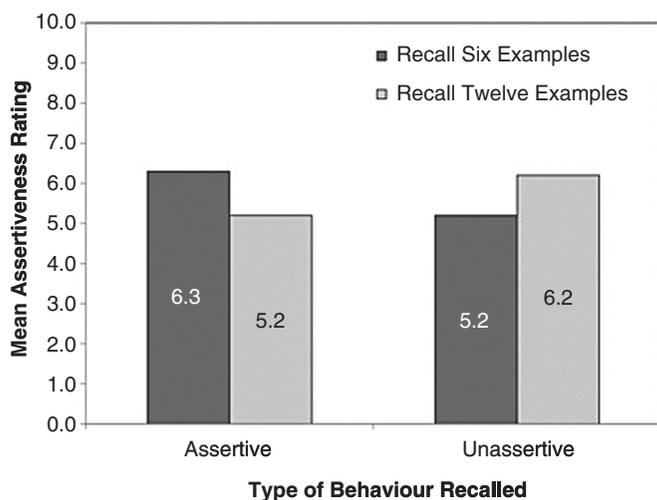
**Figure 8.10** Mean assertiveness rating as a function of number of examples and type of behaviour recalled. High values represent higher assertiveness ratings. Data are from Schwarz et al. (1991, Experiment 1).

unassertive and 10 equalled assertive). As Figure 8.10 shows, subjects who experienced the difficulty of retrieving 12 examples of assertiveness rated themselves as less assertive relative to those who had to recall 6 examples (and vice versa for the subjects who were recalling examples of unassertiveness). As this example illustrates, judgements can be influenced by the availability of relevant information, but that influence depends on the fluency with which that information is retrieved.

## Implicit theories of stability and change

The fluency with which memories are retrieved also plays an important role in how we construct our personal histories. As this chapter has undoubtedly convinced you by now, human memory is a constructive process, and the way in which we remember the narrative of our lives is no exception. Much like the availability heuristic, our knowledge about our current self can bias our judgements about our past self. As Michael Ross (1989) has argued, people possess implicit theories about stability and change, and they use these implicit theories to construct their personal histories. If people believe that they have been consistent over time with respect to a certain attribute (e.g. extraversion) or attitude (e.g. stance on abortion), then they will consider current status and then judge their past self to be similar. Alternatively, if people believe that they have changed, then they will judge their past self to be different. Much of the time people's implicit theories of stability or change are correct, and they can more or less accurately recall their past self. However, when their implicit theories are wrong it can lead to biases in recall and judgement.

A study by McFarland and Ross (1987) provides a prime example of how people's implicit theories of stability can lead them to overestimate the similarity between their present and past selves. In an initial session, they had undergraduate students with steady dating partners rate themselves and their partner on the expected stability of 25 traits (e.g. honesty, intelligence, reliability) over the next two months. In a follow-up session two months later, the students re-rated themselves and their partners on the same traits and

tried to recall their earlier ratings. Students whose ratings were more positive relative to their earlier rating tended to recall their earlier rating as more positive than it was, and vice versa for students whose ratings became more negative. Presumably, the bias that the students exhibited in remembering their past impressions resulted from their implicit theory of stability and the ease with which they could retrieve their current impressions of themselves and their partner.

Ross and Wilson have gone on to explore the related idea of how people's implicit theory of stability or change affects the 'subjective temporal distance' of events (for a brief review see Ross and Wilson, 2003). Subjective temporal distance is a measure of how far away people feel from past selves. In general, people view their present self more favourably than their past self, and they judge their own self-improvement over time to be greater than that of their peers (e.g. Wilson and Ross, 2001). When people evaluate past selves, they often exhibit a bias in their judgement of the distance of past events in which they judge successes to be closer in time and failures to be further in past (e.g. Ross and Wilson, 2002). Interestingly, subjective temporal distance can be manipulated by making events seem more recent or distant, and people tend to be more critical of their former self when an event seems more distant (Ross and Wilson, 2003).

## Future challenges and questions

This chapter has covered paradoxes of behaviour that result from remembering and knowing, task difficulty, interference and the fluency of cognitive processing. We have described 12 paradoxes that all follow the same general form: people behave in a certain way, but their judgements about their behaviour are somehow wrong-headed. They tell an erroneous story about why they behaved as they did, or they predict the opposite of what will happen (e.g. students think repeated reading will produce better recall later than practice in retrieving, yet the recall results later show exactly the opposite). The purpose of our chapter is to call these puzzles to the attention of researchers.

A target for future research, which would make good on the promise implicit in the title of the book, is to discover the neural basis for these puzzles and paradoxes. We did not attempt to describe the neural bases of the illusions and puzzles that we reviewed, but this omission is only partly our fault. If we had tried to link some of these behavioural findings to brain research, our efforts would have been futile because there is a paucity of research in the neuroscience literature about these topics. Although cognitive neuroscientists are starting to look for answers to some of these questions (e.g. Kuhl *et al.*, 2007; Mobbs *et al.*, 2009; Trepel *et al.*, 2005), a more concerted effort is surely needed.

## References

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **20**: 1063–87.

Bahrick, H. P., Bahrick, L. E., Bahrick, A. S., & Bahrick, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science*, **4**: 316–21.

Balota, D. A., Cortese, M. J., Duchek, J. M., *et al.* (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimers type. *Cognitive Neuropsychology*, **16**: 361–84.

Bartlett, F. C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge: Cambridge University Press.

Bäuml, K. H. (2008). Inhibitory processes. In: Roediger, H. L. (Ed.). *Cognitive Psychology of Memory*. Oxford: Elsevier, 195–220.

Bjork, R. A. (1975). Retrieval as a memory modifier. In: Solso, S. (Ed.). *Information Processing and Cognition: The Loyola Symposium*. Hillsdale, NJ: Lawrence Erlbaum Associates, 123–44.

Bjork, R. A. (1994a). Memory and metamemory considerations in the training of human beings. In: Metcalfe, J. & Shimamura, A. (Eds.). *Metacognition: Knowing About Knowing*. Cambridge, MA: MIT Press, 185–205.

Bjork, R. A. (1994b). Institutional impediments to effective training. In: Druckman, D. and Bjork, R. A. (Eds.). *Learning, Remembering, Believing: Enhancing Human Performance*. Washington, DC: National Academy Press, 295–306.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In: Healy, A., Kosslyn, S. & Shiffrin, R. (Eds.). *From Learning Processes to Cognitive Processes: Essays in Honor of William K. Estes, Volume. 2*. Hillsdale, NJ: Erlbaum, 35–67.

Bower, G. H., & Reitman, J. S. (1972). Mnemonic elaboration in multilist learning. *Journal of Verbal Learning & Verbal Behaviour*, **11**: 478–85.

Brown, A. S., & Halliday, H. E. (1991). Crytomnesia and source memory difficulties. *American Journal of Psychology*, **104**: 475–90.

Brown, A. S., & Murphy, D. R. (1989). Cryptomnesia: delineating inadvertent plagiarism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**: 432–42.

Callender, A. A., & McDaniel, M. A. (2009). The limited benefits of rereading educational texts. *Contemporary Educational Psychology*, **34**: 30–41.

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition*, **20**: 632–42.

Carroll, J. S. (1978). The effect of imagining an event on expectations for the event: an interpretation in terms of the availability heuristic. *Journal of Experimental Social Psychology*, **14**: 88–96.

Castel, A. D., McCabe, D. P., & Roediger, H. L., & Heitman, J. (2007). The dark side of

expertise: domain specific memory errors. *Psychological Science*, **18**: 3–5.

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H., (2008). Spacing effect in learning: a temporal ridgeline of optimal retention. *Psychological Science*, **19**: 1095–102.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: a review and quantitative synthesis. *Psychological Bulletin*, **132**: 354–80.

Christina, R. W., & Bjork, R. A. (1991). Optimizing long-term retention and transfer. In: Druckman, D. & Bjork, R. A. (Eds.). *In The Mind's Eye: Enhancing Human Performance*. Washington, DC: National Academy Press, 23–56.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, **58**: 17–22.

Dempster, F. N. (1989). Spacing effects and their implications for theory and practice. *Educational Psychology Review*, **1**: 309–30.

Dunning, D. (2005). *Self-insight: Roadblocks and Detours on the Path to Knowing Thyself*. New York, NY: Psychology Press.

Ebbinghaus, H. (1967). *Memory: A Contribution to Experimental Psychology* (H. A. Ruger & C. E. Bussenius, Trans.). New York, NY: Dover. (Original work published 1885.)

Fischhoff, B. (1975). Hindsight is not equal to foresight: the effect of outcome knowledge on judgement under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, **1**: 288–99.

Fischhoff, B., & Beyth, R. (1975). 'I knew it would happen': remembered probabilities of once-future things. *Organizational Behaviour & Human Performance*, **13**: 1–16.

Foster, J. B., & Magdoff, F. (2009). *The Great Financial Crisis: Causes and Consequences*. New York, NY: Monthly Review Press.

Gallo, D. A. (2006). *Associative Illusions of Memory*. New York, NY: Psychology Press.

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, **16**: 309–13.

Gilovich, T. (1991). *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life*. New York, NY: The Free Press.

Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior*, **15**: 1–16.

Glover, J. A. (1989). The 'testing' phenomenon: not gone but nearly forgotten. *Journal of Educational Psychology*, **81**: 392–9.

Goff, L. M., & Roediger, H. L. (1998). Imagination inflation for action events: repeated imaginings lead to illusory recollections. *Memory & Cognition*, **26**: 20–33.

Goode, M. K., Geraci, L., & Roediger, H. L., III (2008). Superiority of variable to repeated practice in transfer on anagram solution. *Psychonomic Bulletin & Review*, **15**: 662–6.

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

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, **88**: 67–85.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**: 3–28.

Johnson, M. K., Raye, C. L., Wang, A. Y., & Taylor, T. H. (1979). Fact and fantasy: the roles of accuracy and variability in confusing imaginations with perceptual experiences. *Journal of Experimental Psychology: Human Learning and Memory*, **5**: 229–40.

Kahneman, D. (2003). Maps of bounded rationality: psychology for behavioral economics. *American Economic Review*, **93**: 1449–75.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, **15**: 966–8.

Karpicke, J. D., Butler, A. C., & Roediger, H. L. (2009). Metacognitive strategies in student learning: do students practice retrieval when they study on their own? *Memory*, **17**: 471–9.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken as knowing: ease of generation as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language*, **32**: 1–24.

Kerr, R., & Booth, B. (1978). Specific and varied practice of a motor skill. *Perceptual and Motor Skills*, **46**: 395–401.

Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, **14**: 219–24.

Kuhl, B. A., Dudukovic, N. M., Kahn, I., & Wagner, A. D. (2007). Decreased demands on cognitive control reveal the neural processing benefits of forgetting. *Nature Neuroscience*, **10**: 908–14.

Loftus, E., Miller, D., & Burns, H. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, **4**: 19–31.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automation destruction: an example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, **13**: 585–9.

Magill, R. A., & Hall, K. G. (1990). A review of the contextual inference effect in motor skill acquisition. *Human Movement Science*, **9**: 241–89.

Marsh, R. L., & Bower, G. H. (1993). Eliciting cryptomnesia: unconscious plagiarism in a puzzle task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **19**: 673–88.

Marsh, R. L., Landau, J. D., & Hicks, J. L. (1997). Contributions of inadequate source monitoring to unconscious plagiarism during idea generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **23**: 886–97.

Mawhinney, V. T., Bostow, D. E., Laws, D. R., Blumenfeld, G. J., & Hopkins, B. L. (1971). A comparison of students studying-behavior produced by daily, weekly, and three-week testing schedules. *Journal of Applied Behavior Analysis*, **4**: 257–64.

McCabe, D. P., Smith, A. D., & Parks, C. M. (2007). Inadvertent plagiarism in young and older adults: the role of working memory capacity in reducing memory errors. *Memory & Cognition*, **35**: 231–41.

McFarland, C., & Ross, M. (1987). The relation between current impressions and memories of self and dating partners. *Personality and Social Psychology Bulletin*, **12**: 228–38.

Meade, M. L., & Roediger, H. L. (2002). Explorations in the social contagion of memory. *Memory & Cognition*, **30**: 995–1009.

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, **9**: 596–606.

Michael, J. (1991). A behavioral perspective on college teaching. *The Behavior Analyst*, **14**: 229–39.

Mobbs, D., Hassabis, D., Seymour, B., *et al.* (2009). Choking on the money: reward-based performance decrements are associated with midbrain activity. *Psychological Science*, **20**: 955–62.

Newell, A. (1973). You can't play 20 questions with nature and win: projective comments on the papers of this symposium. In: Chase, W. G. (Ed.). *Visual Information Processing*. New York, NY: Academic Press, 283–308.

Newton, L. (1990). Overcofidence in the communication of intent: heard and unheard melodies. Unpublished doctoral dissertation, Department of Psychology, Stanford University.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review*, **84**: 231–59.

Oppenheimer, D. M. (2006). Consequences of erudite vernacular utilized irrespective of necessity: problems with using long words needlessly. *Applied Cognitive Psychology*, **20**: 139–56.

Perfect, T. J., & Stark, L. (2008). Why do I always have the best ideas? The role of idea quality in unconscious plagiarism. *Memory*, **16**: 386–94.

Piaget, J. (1962). *Play, Dreams and Imitation in Childhood*. New York, NY: Norton.

Pohl, R. F. (2004). *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory*. Hove: Psychology Press.

Preston, J., & Wegner, D. M. (2007). The eureka error: inadvertent plagiarism by misattributions of effort. *Journal of Personality and Social Psychology*, **92**: 575–84.

Rajaram, S. (1993). Remembering and knowing: two means of access to the personal past. *Memory & Cognition*, **21**: 89–102.

Roediger, H. L. (1974). Inhibiting effects of recall. *Memory & Cognition*, **2**: 261–9.

Roediger, H. L. (1978). Recall as a self-limiting process. *Memory & Cognition*, **6**: 54–63.

Roediger, H. L. (1980). Memory metaphors in cognitive psychology. *Memory & Cognition*, **8**: 231–46.

Roediger, H. L. (1990). Implicit memory: retention without remembering. *American Psychologist*, **45**: 1043–56.

Roediger, H. L., & Karpicke, J. D. (2006a). The power of testing memory: basic research and implications for educational practice. *Perspectives on Psychological Science*, **1**: 181–210.

Roediger, H. L., & Karpicke, J. D. (2006b). Test-enhanced learning: taking memory tests improves long-term retention. *Psychological Science*, **17**: 249–55.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: remembering words that were not presented in lists. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **21**: 803–14.

Roediger, H. L., Balota, D. A., & Watson, J. M. (2001a). Spreading activation and the arousal of false memories. In: Roediger, H. L., Nairne, J. S., Neath, I. & Surprenant, A. M. (Eds.). *The Nature of Remembering: Essays in Honor of Robert G. Crowder*. Washington, D.C.: American Psychological Association Press, 95–115.

Roediger, H. L., Jacoby, D., & McDermott, K. B. (1996). Misinformation effects in recall: creating false memories through repeated retrieval. *Journal of Memory and Language*, **35**: 300–18.

Roediger, H. L., Meade, M. L., & Bergman, E. (2001b). Social contagion of memory. *Psychonomic Bulletin & Review*, **8**: 365–71.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, **96**: 341–57.

Ross, M., & Wilson, A. E. (2002). It feels like yesterday: self-esteem, valence of personal past experiences, and judgments of subjective distance. *Journal of Personality and Social Psychology*, **82**: 792–803.

Ross, M., & Wilson, A. E. (2003). Autobiographical memory and conceptions of self: getting better all the time. *Current Directions in Psychological Science*, **12**: 66–9.

Rundus, D. (1973). Negative effects of using list items as recall cues. *Journal of Verbal Learning & Verbal Behavior*, **12**: 43–50.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: common principles in three paradigms suggest new concepts for training. *Psychological Science*, **3**: 207–17.

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: another look at the availability heuristic. *Journal of Personality and Social Psychology*, **61**: 195–202.

Shapiro, D. C., & Schmidt, R. A. (1982). The schema theory: recent evidence and developmental implications. In: Kelso, J. A. S. & Clark, J. E. (Eds.). *The Development of Movement Control and Co-ordination*. New York, NY: Wiley, 113–50.

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory*, **5**: 179–87.

Shepard, R. N. (1981). Psychological complementarity. In: Kubovy, M. & Pomerantz, J. R. (Eds.). *Perceptual Organization*. Hillsdale, NJ: Lawrence Erlbaum Associates, 279–342.

Slamecka, N. J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology*, **76**: 504–13.

Starfield, B. (2000). Is US health really the best in the world? *Journal of the American Medical Association*, **284**: 483–5.

Thomas, A. K., Bulevich, J. B., & Loftus, E. F. (2003). Exploring the role of repetition and sensory elaboration in the imagination inflation effect. *Memory & Cognition*, **31**: 630–40.

Trepel, C., Fox, C. R., & Poldrack, R. A. (2005). Prospect theory on the brain: toward a cognitive neuroscience of decision under risk. *Cognitive Brain Research*, **23**: 34–50.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist*, **26**: 1–12.

Tulving, E., & Arbuckle, T. Y. (1963). Sources of intratrial interference in paired-associate learning. *Journal of Verbal Learning & Verbal Behavior*, **1**: 321–34.

Tversky, A., & Kahneman, D. (1973). Availability: a heuristic for judging frequency and probability. *Cognitive Psychology*, **5**: 207–32.

Tversky, A., & Kahneman, D. (1974). Judgments under uncertainty: heuristics and biases. *Science*, **185**: 1124–31.

Underwood, B. J. (1961). Ten years of massed practice on distributed practice. *Psychological Review*, **68**: 229–47.

Wilson, A. E., & Ross, M. (2001). From chump to champ: people's appraisals of their earlier and current selves. *Journal of Personality and Social Psychology*, **80**: 572–84.

Wilson, T. D. (2002). *Strangers to Ourselves: Discovering the Adaptive Unconscious*. Cambridge, MA: Harvard University Press.

Wood, G. (1978). The knew-it-all-along effect. *Journal of Experimental Psychology: Human Perception and Performance*, **4**: 345–53.