

The Role of Repeated Retrieval in Shaping Collective Memory

Henry L. Roediger, III, Franklin M. Zaromb, & Andrew C. Butler

This book is about culture and memory, about how the society and culture in which people grow up helps to determine their individual memories, collective memories, and identity. We will emphasize how the process of repeated retrieval helps to shape our memories. Before we get to our main story, we need to provide some background.

Many different cultures exist within the community of scholars who study memory, from humanistic approaches to a whole variety of scientific approaches to the field (see Roediger, Dudai & Fitzpatrick, 2007). In our chapter we attempt to blend some insights from several approaches to studying memory. In particular, we apply the principles that have emerged from research in experimental cognitive psychology to issues in collective memory in hopes that combining insights from various disciplines may point the way to progress in understanding larger issues in the study of memory.

Our chapter has six parts. First, we discuss issues in collective memory – how we conceive the topic – by considering three conceptual oppositions (following Wertsch & Roediger, in press). We then discuss how these issues might play out in learning about history from textbooks. Then, we describe three main mechanisms that we believe play a critical role in shaping collective memory: the act of retrieving information from memory; repeated retrieval of information over long periods of time; and finally the role of feedback in modifying memories. The third, fourth, and fifth sections of the chapter spell out these mechanisms and how they might contribute to the creation of collective memories. The sixth section explores the issue of accuracy in collective memory, and how it is affected by the mechanisms of retrieval and feedback. Finally, we end with some conclusions.

What is Collective Memory?

Collective memory is not a term with a precise definition. Since Maurice Halbwachs introduced the term in the 1920s, it has been used in many different ways. Historians, sociologists, anthropologists, psychologists, and literary analysts all have slightly different definitions of what constitutes a collective memory. The general feature that unites all uses of the term is that collective memory is a form of memory

that transcends individuals and is shared by the group. In our treatment in this section, we rely heavily on observations of Dudai (2002) and Wertsch and Roediger (in press).

In *Memory from A to Z*, Yadin Dudai (2002) noted that “The term ‘collective memory’ actually refers to three entities: a body of knowledge, an attribute, and a process” (p.51). The body of knowledge refers to the common knowledge of a society or culture (a European) that can be divided further into subgroups (Italians, Florentines, Florentine women over 60). Of course, one can go further to the level of families, friends, and even pairs of people, where the collective memories are even more specialized. Regardless of the size of the group, each individual in the group will share various memories with other members of that group. Importantly, the collective memory is present in each individual; that is, each member of the group would produce roughly the same information if asked about the memory individually. In this sense, it is different from the concept of transactive memory, in which people serve as external memory aids to one another (Wegner, 1986). The attribute to which Dudai refers is “the distinctive holistic image of the past in the group” (such as World War II veterans in the U.S., to whom some refer as the “greatest generation” of Americans; Brokaw, 2001). Finally, the process is the continual evolution in understanding between the individual and the group. That is, individuals may influence and change the collective memory of the group, and the group can change the individual’s understanding of being a member of the group and of the group’s past.

These three features – body of knowledge, attribute, and process – are critical to understanding collective memory. To flesh out the concept more fully, we follow Wertsch and Roediger (in press) in briefly considering a set of three oppositions that highlight different uses of the term.

Collective memory versus collective remembering. The term collective memory is sometimes used to refer to a relatively static set of knowledge, as in Dudai’s (2002) “body of knowledge” feature or Tulving’s (1972) semantic memory. This might refer to what people in England know about what happened in 1066, or what most Americans know about their Revolutionary War or about the events of 9/11, and so on. On the other hand, collective remembering refers to a dynamic process, how different perspectives and reconstructions over the years can provide a contentious process of representing the past. Examples in U. S. history include how the view of Native Americans (“Indians”) and African Americans has changed over the years, often with the “new” view espoused by each group opposing that of the dominant culture. Such disputes over how some central event is to be remembered erupt during construction of museums and monuments, the writing of textbooks, and creation of national holidays. These disputes over how the past is to be remembered that are fought among

groups with different interests have been termed “the politics of memory” (Kramer, 1996).

History versus collective remembering. Halbwachs (1980, 1992) raised the issue of how collective memory – a representation of the past – might be different from history. In contemporary thinking, history and collective memory are often considered to be conflicting. The reason is that the two modes of knowing have different aspirations of representing the past. History aspires to provide an accurate account of the past, even when that record may reflect poorly on the people being represented. In contrast, in collective remembering, the past is often tied to the present, so that a person’s self identity and group identity are buoyed by the glorious past history of the people. In collective memory, negative aspects of the past may be omitted or suppressed in the service of representing what kind of people we are. Americans today grapple with how to reconcile “America the good” with atrocities committed in the past against Native Americans and African slaves, as well as to many immigrant groups soon after they arrived (not to mention scandals more recently involving abuse of prisoners). As noted by the historian Novick (1999):

“To understand something historically is to be aware of its complexity, to have sufficient detachment to see it from multiple perspectives, to accept the ambiguities, including moral ambiguities, of protagonists’ motives and behavior. Collective memory simplifies; sees events from a single, committed perspective; is impatient with ambiguities of any kind; reduces events to mythic archetypes.” (pp. 3-4)

Individual versus collective remembering. Remembering is often conceived as occurring within an individual: some cue reminds one of events or facts of the past. Bartlett (1932) even raised the issue of whether collective remembering can exist, at least in the strong sense that a group (as a group) “can be usefully characterized as having some sort of memory in its own right” (Wertsch, 2002, p. 22). After all, memories must be carried in individual nervous systems. In contrast to this position, Schudson (1992) has questioned whether memory can be anything but collective in that “memory is social...it is located in institutions rather than in individual human minds in the form of rules, laws, standardized procedures and records, a whole set of cultural practices through which people recognize a debt to the past” (p. 347). These views are quite different, but the point of contact may be “agreement on the point that socially situated individuals are the agents of remembering” (Wertsch & Roediger, in press).

As Wertsch (2002) has pointed out, individuals use a set of cultural tools to aid in representing the past – written symbols in books, paintings and icons, museums and monuments, and today the huge amount of information distributed on the internet.

Most readers doubtless have used Google or Wikipedia to help remember some event or fact. Further, our memories are socially constructed in that we rely on others to provide information to update our memories of events, which can both improve our recollections and can insert errors, depending on the information provided (Roediger, Bergman & Meade, 2001). In short, individuals remember, but with much help from social and cultural aids including reports of memories from other people. Collective memories may ultimately reside in individual nervous systems, but complex social processes provide the basis for our memories and their nature.

The foregoing provides our reflections on the nature of collective memory (see Wertsch, 2002; in press) for admirable overviews of the field. We turn now to specific processes by which we believe collective memories (indeed, all memories) are created and maintained over long periods of time. We refer first to retrieval of memories. Much evidence points to the critical role of retrieval in future remembering (e.g., Roediger & Karpicke, 2006a). Although we do not have evidence to completely back up this point, we strongly suspect that the only events occurring today that can be retrieved far into the future are ones that have been retrieved previously. That is, it is probably impossible if you are over 20 to ever accurately retrieve an event from when you were ten unless you have retrieved it during the interim. We next turn to repeated retrieval and spaced retrieval as keys in maintaining memories over long periods of time, with the theme being a “use it or lose it” point of view. Further, the process of retrieval is itself shaped by feedback given, either from other people or from sources such as history books, novels, television and the like. We consider these issues in turn before continuing on to the issue of truth and myth (accuracies and inaccuracies) in collective memory.

History Textbooks: A Potential Source of Collective Memories

The classroom is a setting in which many collective memories are likely to be created, because people learn much of their historical knowledge through formal schooling. Although students are sometimes taught to consider multiple perspectives on historical events, the historical information presented to students is often closer to the aims of collective remembering than those of history (as discussed above). That is, historical information is often portrayed in ways that conform to present cultural values or political objectives, rather than providing an objective account of the past. Indeed, when it comes to teaching history in educational settings, students primarily learn accounts of the past that reflect the current biases, goals, and beliefs of a particular social group or nation-state (Loewen, 1995; see Blatz & Ross, this volume). This contrast can also be observed in a variety of other contexts in which history is taught, commemorated, or symbolically represented, such as museums and national

holidays (see Blatz & Ross, this volume; Blight, this volume; Pennebaker & Gonzales, this volume; Winter, this volume).

In the classroom, the textbook used by the teacher usually dictates the account of history that is taught and, as a result, textbooks are likely to be a primary source of information for collective memories. Although one might assume that the information about and perspective taken on historical events has remained largely consistent over time, there is actually a high degree of variability. In a recent survey of U.S. history textbooks from 1794-1999, Ward (2006) analyzed how textbooks have portrayed seminal events in U.S. and modern world history over time. History textbooks are especially useful for studying collective memory, because, as Ward argues, they typically reflect the most widely read historical accounts in any country at a given time and are studied during a student's formative adolescent years.

This point is confirmed by a study conducted in our lab in which we asked American college students to recall the 10 most important events that occurred during the U.S. Civil War (1861-1865), World War II (1939-1945), and the current Iraq War (2003-Present). Afterwards, students answered a series of questions about their knowledge and attitudes concerning the specific events they recalled. When asked about the particular source from which they first learned about each of the recalled events, students recollected having first learned 85% of the events listed for the U.S. Civil War from either textbooks (55%) or teachers (30%). Similarly, 77% of the events listed for World War II were first learned from textbooks (48%) or teachers (29%). Of course, given that the Iraq War is too recent to have been taught in the classroom, students instead recollected having first learned of the majority of listed Iraq War events from media sources (87%) such as television (62%), print media (19%), and the internet (6%).

Ward (2006) further argues about the importance of textbooks in shaping our collective knowledge and conceptions of history:

“Unlike independently authored historical accounts, textbooks are a quasi-official story, a sort of state-sanctioned version of history. In nearly all countries the government takes some role in setting the standards for an acceptable cultural, political, and social history—i.e., what the authorities want the next generation to learn about its own national heritage—enfolding them, as it were, into a collective national identity. The fundamental contribution of any textbook is, of course, the story, an account that reveals much about a nation's priorities and values. Because history textbooks contain national narratives written by national authors for a national audience, they model the national identity in a very profound and unique way.” (pp. xviii-xix)

Consider the events surrounding the dropping of the atomic bomb on Hiroshima and Nagasaki in 1945. While there is no doubt as to what actually occurred, over the past 60 years, U.S. history textbooks have described the rationale and significance of these events in ways that appear to reflect prevailing attitudes in American society at different times toward nuclear weapons and the consequences of their use. In 1947, U.S. history textbooks emphasized the role that the atomic bombs played in hastening Japanese surrender, but by 1954, near the height of the Cold War nuclear arms race, textbooks emphasized the development of the bombs themselves, skipping the subject of their strategic role in ending the war (Ward, 2006, pp. 289-292). In the wake of the 1961 Cuban Missile Crisis where the U.S. and the Soviet Union reached the brink of a nuclear war, one textbook from 1966 informed students in its discussion of the 1945 atomic bombings that “a nuclear holocaust might be the end result, if a ‘lasting peace’ was not found” (pp. 291). It was not until the 1990s, in the aftermath of the Cold War and the break-up of the Soviet Union that U.S. textbooks began to provide students with several interpretations as to why the atomic bomb was used at the end of World War II.

Even more striking are the differences in historical accounts revealed by comparing textbooks used around the world. Contemporary Japanese textbooks question the U.S. motives for dropping the atomic bombs and emphasize the devastating impact of the bombs and the cost in human lives. Canadian textbooks highlight the role that Canadians played in mining and enriching the uranium for the atomic bomb. Likewise, British textbooks remind students of the contributions of British and other European scientists to developing the atomic bomb. By contrast, Italian textbooks focus on U.S. motivations by explicitly stating that there was no military justification for using nuclear weapons, and that President Truman was primarily concerned with ensuring a dominant role for the U.S. in post-war international affairs, especially with respect to the Soviet Union (Lindaman & Ward, 2004, pp. 238-244).

Judging by the information that is both emphasized and omitted in history textbooks, it comes as no surprise that students’ knowledge and understanding of history can dramatically vary from generation to generation or from country to country (e.g., Wertsch, 2002; Liu et al., 2005). For example, Wertsch (2002) reported cases that pitted the recollections of World War II of Russian and American students that revealed striking differences in the specific events the two groups recalled as well as the basic structure used to recount the war’s narrative.

Wertsch (2002; see also Wertsch, in press) described one exercise that required participants to list the most important events of World War II. Among American students, the most frequently listed events were: 1) Attack on Pearl Harbor (December 7, 1941); 2) Battle of Midway (June, 1942); 3) D-Day (June 6, 1944); 4) Battle of the

Bulge (Winter 1944-1945); 5) Holocaust (throughout the war); and 6) Atomic bombing of Hiroshima and Nagasaki (August 1945). Data collected in our lab from American college students generally replicated these findings, with the following events listed most frequently: 1) Atomic bombing of Hiroshima and Nagasaki; 2) Attack on Pearl Harbor; 3) D-Day; 4) Holocaust; 5) German invasion of Poland (September, 1939); and 6) U.S. entering World War II. In contrast, Russian students surveyed during the 1990s generated a completely different list of events, which included: 1) German attack on USSR (June 22, 1941); 2) Battle of Moscow (Winter 1941-1942); 3) Battle of Stalingrad (Winter 1942-1943); 4) Battle of Kursk salient (Summer 1943); 4) Siege of Leningrad (1942-1944); and 5) Final Battle for Berlin (1945).

What is most striking about these lists of events is the fact that there is no overlap. Wertsch (in press) explains that the lack of overlap does not imply that participants in either of the two groups have no knowledge of the others' listed items. Rather, the events listed for one group are considered much less important than those listed by the other group. For example, what is listed as "D-Day" by American students is known to Russians as the "opening of the second front" in June of 1944. The first and primary front in their eyes was the Eastern front.

Although group-related differences in historical memory may be explained in terms of what students learn in the classrooms of different countries, it is of equal if not greater importance to consider how students learn history. That is, what educational practices enhance learning of materials taught in the classroom? We now turn to exploring how the mechanisms of retrieval, repeated retrieval (or spacing), and feedback can help to create, maintain, and shape collective memories, respectively. For each of these mechanisms, we will first review some of the research from the laboratory and then discuss how the principles that emerge can be applied to collective memory.

Retrieval: A Mechanism for Creating Collective Memories

One way in which people learn information is through the process of retrieving that information from memory. We now review findings from the laboratory that demonstrate the powerful effect of retrieval on long-term retention. Of course, the mnemonic benefits of retrieval are not confined to the laboratory. For example, retrieval practice is regularly incorporated in the classroom in the form of tests. Accordingly, we also explore how the mechanism of retrieval may play a role in the creation of collective memories.

Laboratory Research on the Testing Effect

A robust finding in the cognitive and educational psychology literature is that testing subjects on their ability to remember previously learned material improves

long-term retention relative to re-studying it for an equivalent amount of time (e.g., Carrier & Pashler, 1992; for review see Roediger & Karpicke, 2006a). This finding, known as the testing effect, occurs even when neither feedback nor further study opportunities are provided (Gates, 1917; Glover, 1989; Roediger & Karpicke, 2006b; Spitzer, 1939). Interestingly, the benefits of testing have also been shown to promote retention of related, but un-tested information (Chan, McDermott, & Roediger, 2006). More importantly, the benefits of testing have been shown to be quite powerful, both in the laboratory and the classroom (for a meta-analysis see Bangert-Drowns, Kulik, & Kulik, 1991). For example, Jones (1923-1924) reported that when students were tested after a classroom lecture, they scored twice as high as untested students on a retention test given eight weeks later.

In a recent review of the steadily growing body of research on the testing effect, Roediger and Karpicke (2006a) argued that testing has a direct, positive effect on learning, as well as many indirect benefits. The direct effect of testing derives from the use of retrieval as a mechanism to enhance the long-term retention of information. Retrieving information from memory is not a neutral event; rather, it leads to a modification of the memory trace (Bjork, 1975). As a result, the act of retrieving information from memory produces a memory trace that is more resistant to forgetting. The indirect benefits of testing might include helping students to direct the focus of their subsequent studying strategies (i.e. “learning to learn”), encouraging students to study more frequently, and reducing students’ anxiety about taking tests because their grade is not completely dependent on a small number of tests (e.g., a mid-term and final) and they are familiar with the types of questions that might appear on the test.

Although most of the testing effect studies reported in the research literature have been obtained in laboratory settings using discrete verbal study materials (e.g., lists of words), several classic studies conducted with relatively large populations of grade-school students have also shown strong, positive effects of testing (Gates, 1917; Jones, 1923-1924; Spitzer, 1939). Recent years have also seen a flurry of studies that have investigated the testing effect using more educationally-relevant materials, such as prose passages and textbooks (e.g., Foos & Fisher, 1988; Kang, McDermott, & Roediger, 2007; McDaniel et al., 2007; Roediger & Karpicke, 2006b).

Of relevance to history education, Butler and Roediger (2007) examined the benefits of testing on long-term retention of history lectures in a simulated classroom setting. Subjects viewed a series of lectures on consecutive days and, following each lecture, engaged in one of three types of activities: 1) studying a lecture summary; 2) taking a multiple-choice test; or 3) taking a short answer test. One month later, students completed a final short answer test. Not surprisingly, studying the lecture summaries or taking a multiple-choice test led to better performance on the final

recall test than a no activity control condition. More importantly, taking an initial short-answer test significantly improved final recall relative to studying the lecture summary or taking an initial multiple-choice test. The superiority of the short answer testing condition compared to the study and no activity conditions indicates a direct long-term benefit of testing. However, the advantage of short answer testing over multiple-choice testing suggests that test format, and more specifically, the cognitive demands of particular tests on retrieval processes are critical in determining testing effects (Kang, McDermott, & Roediger, 2007). Whereas students may rely upon the familiarity of responses to answer multiple-choice test questions, they must engage in effortful retrieval in order to complete short-answer recall tests.

Retrieval Creates Collective Memories

Although it has not been empirically investigated, one might reasonably argue that tests play an additional role of designating the historical information and perspectives that are important to know (as well as the information that is of secondary or little importance). This is especially true of standardized or AP exams that are administered at the state or national level. For instance, American students are much more likely to be tested on their knowledge of the Japanese attack on Pearl Harbor than the German siege on Leningrad, even though both events were significant in their own right in determining the course of World War II. In other words, tests may determine the importance of particular historical events and concepts in the minds of students.

Returning to the comparison of American and Russian students' recollections of World War II, Wertsch (2002) noticed distinct patterns in the basic narrative structure of their recall testimonies. He argued that these patterns reflect the use of what he terms a "schematic narrative template" to recount historical episodes. Similar to Bartlett's (1932) conception of schemata, schematic narrative templates are abstract knowledge representations that are readily accessible to the narrator and guide the reconstructive process of remembering. As an example, Wertsch (2002) proposed a schematic narrative template for Russians entitled the "Triumph over Alien Forces" narrative. It may be distilled from accounts of events that occurred throughout Russian history such as Napoleon's invasion of Russia in the early 1800s and the German invasion of World War II. The basic elements of the schematic narrative template are: 1) Russia is peaceful and not interfering with others; 2) Russia is viciously and wantonly attacked without provocation; 3) Russia almost loses everything in total defeat; and 4) Through heroism and exceptional bravery, and against all odds, Russia triumphs.

To the extent that tests either reflect or emphasize the importance of particular historical facts, they force students to actively retrieve historical knowledge and attempt to relate that information to the larger scope of national or world history as

well as to the present. As such, tests might even serve a function that is not unlike commemorative holidays such as President's Day, the Fourth of July, Veteran's Day, or Martin Luther King, Jr.'s Birthday, because, in essence, national holidays stress the role of particular historical events and figures in shaping national history. As days of remembrance, national holidays are designed to encourage citizens to recollect, discuss, and ponder the significance of past wars, leaders, national tragedies, or the birth of political institutions. Together, the testing of history in the classroom and the yearly remembrance of key historical events and figures can serve as important tools for promoting the long-term retention of historical knowledge and the shaping of a collective memory.

Repeated Retrieval and Strengthening Of Collective Memories

If retrieval is a mechanism that creates collective memories, repeated retrieval is the process through which these collective memories are strengthened and maintained over long periods of time. The repeated retrieval of information from memory has been shown to produce better long-term retention than a single retrieval attempt (e.g., Wheeler & Roediger, 1992). In addition, when the repeated retrieval attempts are spaced over a period of time, much greater retention is produced than when the attempts are massed together; this finding, commonly referred as the spacing effect, is one of the oldest in the literature, dating back to the first empirical studies of memory (Ebbinghaus 1885/1967; see Roediger, 1985). We now review some of the findings from the laboratory on repeated retrieval and spacing, and then attempt to apply the idea to the long-term maintenance of collective memories.

Laboratory Research on Repeated Retrieval and Spacing

Cognitive psychologists have investigated how the timing and spacing of repeated retrieval influences long-term retention using the experimental paradigm of paired-associated learning. Specifically, if subjects learn a pair of associates (A-B), where A might be "horse" and B its Spanish equivalent "caballo", is it best to practice retrieving the A-B pair in massed fashion with retrieval attempts occurring in immediate succession, or is it better to space the retrieval attempts over intervals of time? At first blush, massed retrieval practice is appealing, because it occurs shortly after initial learning when little forgetting has occurred, and as a result, retrieval may be errorless. On the other hand, spaced retrieval may be superior, just as repeated study benefits long-term retention when the presentation of the to-be-learned items is repeated in spaced as opposed to massed fashion (Glenberg, 1976; Melton, 1970; for a review, see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Dempster, 1989). Moreover, if spaced retrieval is beneficial, should individuals practice retrieving information using the

same, equal interval or, rather, should the interval vary, such as in an expanding schedule?

Landauer and Bjork (1978) conducted two experiments that compared these schedules of repeated retrieval. In their experiments, subjects studied pairs of items (e.g., names paired with surnames or names paired with faces) and then took a series of cued recall tests (without feedback) for each of the study pairs distributed according to massed or spaced (equal and expanding interval) schedules in a continuous paired-associates learning task. Subjects then completed a final recall test 30 minutes after the learning phase. They found that the equal-interval retrieval schedule produced better final recall test performance than did massed retrieval. A number of subsequent studies have replicated this long-term memorial advantage of spaced versus massed retrieval practice (e.g., Cull, 2000; Cull, Shaughnessy, & Zechmeister, 1996). In addition, the expanding-interval schedule produced better final recall test performance than the equal-interval schedule. However, many subsequent studies have failed to demonstrate an advantage of one type of spacing schedule over the other (see Balota, Duchek, and Logan, 2007).

Recent work has sought to break the impasse by redirecting the focus of attention to the timing of the initial test, or retrieval attempt (e.g., Karpicke & Roediger, 2007a; Logan & Balota, 2007). These studies have shown that the timing of the initial recall test in a spaced, repeated retrieval schedule may enhance or detract from the success of the schedule. If the first test occurs immediately after learning, as is usually the case in expanding retrieval schedules, then the memorial benefit of such a test is minimal, because little retrieval effort is required to successfully recall the just-presented information. However, if the initial test occurs after a delay (provided it is not too long) as is customary with equal-interval schedules, then the first test will require some retrieval effort, thereby promoting long-term retention.

Repeated Retrieval Maintains Collective Memories

To the extent that repeated retrieval does play an important role in the formation and shaping of collective memory, laboratory studies of repeated retrieval and spacing have important implications for remembering historical events. First, the spacing effect literature underscores the importance of considering the different times in which historical knowledge is learned and remembered in society. For instance, many historical facts taught in schools at the primary and secondary level are learned and tested in massed fashion and may not be encountered again for many years. On the other hand, holidays such as the Fourth of July encourage individuals to recollect the events surrounding the signing of the Declaration of Independence at equally spaced yearly intervals. Other remembrance occasions such as school reunions may first take place shortly after schooling (e.g., 5 years), then continue to be held at expanding

intervals (e.g., 10 years, 25 years, 50 years). Given the variety of schedules, it is worth asking how the timing with which individuals repeatedly think about or discuss their recollections of a particular historical event in relation to when it was first learned or experienced influence their memory for the event. To address this question, we will discuss several studies of the “flashbulb memory” phenomenon, because they provide some evidence for the influence of repeated retrieval, as well as the frequency and timing of retrieval, on the formation and long-term maintenance of collective memories.

While the educational practices and collective rituals of social groups play a significant role in the development and shaping of collective memory, they do not have absolute control over the process. This is because many events that have greatly affected the course of history have occurred unexpectedly. Such events might include the stock market crash of 1929, the assassination of John F. Kennedy in the U.S. or Yitzhak Rabin in Israel, or the terrorist attacks of 9/11. As revealed by studies of the “flashbulb memory” phenomenon, the psychological impact of shocking public historical events can have long-term memorial consequences, not only for individuals, but for larger social groups as well.

Flashbulb memories are memories for specific events that are so vivid and personal that an individual seems to remember exact details of the context in which he or she first learned or experienced the event for a long period of time (Brown & Kulik, 1977). As Brown and Kulik (1977) reported in their pioneering study, individuals can develop flashbulb memories in response to any event that is very surprising or shocking and of personal significance, such as learning about the death of the close relative. However, in many cases, flashbulb memories develop in a large number of individuals in response to shocking public events, such as the assassination or attempted assassination of a political leader. For young adult participants in the Brown and Kulik study in the late 1970s, the assassination of President John F. Kennedy in 1963 was such an event and to this day serves as the best known example of the flashbulb memory phenomenon. Over the past 30 years, researchers have debated the existence, accuracy, and methods for assessing flashbulb memories (e.g., Conway, 1995; Neisser & Harsch, 1992; Rubin & Kozin, 1984). Nevertheless, for our purposes, it is worth highlighting findings that elucidate the role that retrieval processes may play in the formation and shaping of memories for historical events such as 9/11 among social groups.

Less well known is the fact that Brown and Kulik (1977) examined whether flashbulb memories tend to arise differently in distinct groups of people. The authors administered questionnaires to White and African-American participants that asked whether they remembered events associated with various contemporary political leaders. Whereas some leaders were political officials in government, others were

leaders in the Civil Rights movement, and while most of the listed events were assassinations or assassination attempts, one event involved an accidental drowning (Mary Jo Kopeckne, riding in Ted Kennedy's car) and another was a death due to natural causes (General Francisco Franco of Spain). Participants were also asked to remember a personal, unexpected shock, such as the death of a friend or relative, serious accident, etc. If they remembered a particular event, they composed a detailed narrative describing the circumstances in which they learned of the event. Last, they were asked to rate the importance or consequentiality of the event and the number of times they repeated telling others about the event.

Brown and Kulik (1977) found that events that were quite surprising and highly consequential tended to elicit flashbulb memories. Such events were also more likely to be retold numerous times to others. More importantly, there was a greater tendency for African American participants to report flashbulb memories for the assassinations of Medger Evers, Malcolm X, Martin Luther King Jr., and George Wallace than White participants.

Following this initial study, a number of recent studies have also reported similar types of group differences in memory for historical events. For instance, Berntsen and Thomsen (2005) examined Danes' memories for the German invasion of Denmark in 1940 and its liberation in 1945. The study investigated differences between the memory of individuals who had ties to the resistance movement and those who did not. Berntsen and Thomsen (2005) found that participants who reported ties to the Danish resistance movement had more vivid and accurate memories than the other participants. Danes with ties to the resistance movement also reported ruminating on and discussing these events with other people more than Danes not involved in the resistance movement. Repeated retrieval probably helped consolidate the vivid memories.

Curci and Luminet (2006) examined influence of national identity on the accuracy, durability, and subjective experiences associated with memories for the terrorist attacks of September 11, 2001 (see also Luminet et al., 2004). Participants in six countries (France, Italy, the Netherlands, Romania, Japan, and the U.S.) were asked to complete questionnaires that assessed their memories for the September 11th attacks. They answered questions about the context in which they first learned of the attacks and their knowledge of the event itself. In addition, they rated the novelty, surprise, importance or consequentiality of the attacks, as well their emotional reactions and background knowledge. In order to examine the consistency and retention of subjects' flashbulb memories, questionnaires were administered a few weeks after the attacks and again about a year later.

As expected, Americans' recollections of the September 11th attacks were more accurate and displayed more of the specific attributes associated with flashbulb

memories than people of other nationalities. In addition, Americans reported stronger emotional responses to the event, more background knowledge, and an increased tendency to ruminate (with the exception of Dutch participants) than other nationalities. Interestingly, the pattern of results for Americans was similar to those for the Italian and Dutch participants, who displayed high specificity and consistency in their flashbulb memories and whose event recollections were generally accurate. All three groups reported strong emotional reactions upon learning of the attacks and frequently rehearsing news about the event by sharing it with others, ruminating, and following relevant media coverage. By contrast, the remaining participants from Belgium, Romania, and Japan, while still having reported highly specific and consistent flashbulb memories of the September 11th attacks, were much less accurate in their recollections of the events themselves.

These results highlight three group-related differences in historical memory. First, although previous studies of flashbulb memories for earthquakes have shown that physical proximity to the event influenced participants' recollections (Er, 2003; Neisser, Winograd, Bergman, Schreiber, Palmer, & Weldon, 1996), the findings of Curci and Luminet (2006) were only partially consistent with this finding. Although Americans reported the most specific, consistent, and accurate memories, flashbulb memories were high and consistent across all national groups. Second, some of the idiosyncratic results are difficult to interpret without considering culturally-specific explanations. For example, Japanese participants reported having the weakest emotional response to the September 11th attacks. Romanian participants were least likely to share news of the attacks with others and to follow media coverage of the event, perhaps reflecting the enduring cultural legacy of past dictatorships. Third, there were significant national differences in event recall accuracy.

It should be noted that at the outset Brown and Kulik (1977) offered little justification for comparing the recollections of White and African Americans except for making the assumption that differences might be observed between individuals in any two social, professional or racial groups; their intuitions led them to predict that some of the recent assassinations of political leaders involved in the Civil Rights movement would resonate more among African-Americans than among Whites (see footnote, p. 77). Brown and Kulik offered no further explanation as to what specific aspects of group identification were associated with group differences in historical memory. However, they did argue that a surprising or highly consequential event is of biological significance and therefore may activate neural mechanisms that enhance the encoding of contextual information related to the circumstances in which the event was originally experienced. To the extent that individuals in groups differ in their interests, different events will seem more surprising or consequential depending on group membership. Furthermore, the psychological impact of such events may lead individuals to rehearse their memories for those experiences in both covert (i.e.,

rumination) and overt manners (i.e., by writing or verbally discussing recollections with others), which in turn, may directly influence how elaborate a person's recollections are of the original event.

Although one might argue along the lines of Brown and Kulik (1977) that the observed group differences in historical memory were primarily determined by the memorial encoding conditions—that is, the initial impact of the event and the influence of group identification on that experience. Alternatively, we argue that of equal if not greater importance may be the role that retrieval plays in the formation and retention of memories across groups over time. For instance, in the above-mentioned studies, memory for historical events was enhanced for groups that tended to retrieve covertly (to ruminate over) and to recollect those events in social contexts (overt retrieval). Probably the repeated retrieval of events is critical to their long-term retention, as has been shown in laboratory studies by Karpicke and Roediger (2007b).

The Timing and Effects of Spaced Retrieval

Shapiro (2006) has offered some further insight into the effects of retrieval by examining college students' recollections of the terrorist attacks on September 11th, 2001, over a two-year period of time. Specifically, she asked whether flashbulb memory (memory for reception context) and event recall were differentially affected by retention interval. In addition, she investigated whether the frequency and temporal spacing of repeated recollections moderated the effects of retention interval on recall. In one experiment, three groups were tested either one or three times for their recollections of 9/11. Three recall tests were completed by the first two groups at either expanding or equally-spaced intervals. In the expanding interval group, subjects were tested 1-2 days, two weeks, and 11 weeks after the attacks of 9/11. The equal spacing group was tested 1-2 days, 6 weeks, and 11 weeks following 9/11. The third group served as a control and was tested only one time, 11 weeks after 9/11.

The memory tests were designed to determine whether subjects developed flashbulb memories for the events surrounding 9/11, as well as assess their knowledge of the events themselves. Based upon the original criteria proposed by Brown and Kulik (1977), flashbulb memory reports were based on subjects' ability to report their location when they first learned of the 9/11 attacks; what activity they were doing at the time, the source of the information from which they first learned about the attacks, and any personal relevance. For the first two groups, the same queries were made during the follow-up assessments in order to evaluate the consistency of their recollections over time. To assess knowledge of the events themselves, subjects were instructed to write down everything they knew about the 9/11 attacks.

Shapiro (2006) found that the length of the retention interval did not affect recall elaboration or consistency for details of the reception context of 9/11. By contrast,

recall for various aspects of the events pertaining to 9/11 was detrimentally affected by a long retention interval. Long-term retention of the events of 9/11 was enhanced by covert retrieval, or rehearsal, as indicated by elevated recall performance and degree of elaboration in the groups assessed three times. However, somewhat surprisingly, the spacing of the assessments did not significantly affect event recall, which suggests that it was the frequency, and not the temporal spacing, of the retrieval attempts that mattered.

In a follow-up experiment, Shapiro (2006) examined the effects of using longer retention intervals of 23 weeks, 1 year, and 2 years on memory for the reception context and events of 9/11. Similar to the results of her first experiment, neither retention interval nor the number of retrieval attempts affected the amount and consistency of reception context information recalled. By contrast, event recall performance, but not consistency, was detrimentally affected by the longer retention intervals. However, even following the longer retention intervals, event recall showed improvement with subsequent retrieval attempts. Using additional questionnaires to probe the subjective experiences associated with subjects' recollections, Shapiro observed that recall was enhanced for subjects that reported their memories as vivid, frozen, or encompassing a longer period of time. Furthermore, recall performance was positively correlated with measures of confidence in accuracy and reported frequency of rehearsal through discussion. The latter provides further support for the idea that repeated overt retrieval aids maintenance of flashbulb memories.

Taken together, these findings are consistent with the notion that the processes underlying the formation of flashbulb memories are different from those involved in the learning and retention of historical event information. Whereas variations in retention interval and retrieval schedules had little impact on flashbulb memories, these variables significantly influenced recollections of what events had actually occurred on 9/11. These findings lend indirect support to Brown and Kulik's (1977) original theory of flashbulb memories, in which encoding mechanisms play a determining role in the formation of memories, but then repeated retrieval helps to maintain memories for the event itself, but not for the surrounding context. If, as Shapiro (2006) observed, repeated retrieval has little influence on one's memory for reception context, then the primary determinants of flashbulb memories must interact at the encoding stage.

By contrasting the durability of flashbulb memories with the fragility of historical event memory, one can appreciate the paradox that societies face when coming to terms with national tragedies such as 9/11. On the one hand, certain subjective features of the event are seemingly unforgettable, while core information is readily forgotten, unless it is rehearsed on a frequent basis. Of course, only the single study by

Shapiro (2006) has yet tackled these issues, so further evidence on these points would be welcome.

Feedback: A Mechanism for Shaping Collective Memories

Feedback consists of information provided or generated after recollection of an event (or the performance of a behavior) that allows the learner to improve future recollection or performance. Feedback can be provided by either an external or internal source. With such a broad definition, the term feedback applies to many different phenomena: from signals generated by the somatosensory system after fine motor movements to comments made by audience members after hearing a presentation. For the purpose of the present chapter, we focus on feedback that is given in the context of acquiring and using declarative knowledge. More specifically, feedback is conceptualized as information about the accuracy of a person's performance on a task that requires retrieval from memory (e.g., recalling the name of the 16th President of the United States on a test, remembering one's 18th birthday, etc.). We first review research from the laboratory on the effect of post-test feedback on learning and retention. Then, we describe how the same feedback mechanisms that influence learning in the laboratory may also shape collective memories outside the laboratory.

Laboratory Research on Feedback

Laboratory studies have investigated a multitude of different factors that influence the effectiveness of feedback after a test (for reviews see Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Butler & Winne, 1995; Hattie & Timperley, 2007; Kulhavy & Stock, 1989), including the timing of feedback, the type of to-be-learned material, the format of the test, the retention interval, and the amount of prior learning (to name but a few). Of these many factors, the most important is the nature of the information provided. In its most basic form, the feedback message must provide learners with information about the veracity of the response (i.e. whether it is correct or incorrect). When additional content is included, the basic feedback message can be expanded along three dimensions: type, form, and load (Kulhavy & Stock, 1989). The different types of feedback are generally classified as a function of the type of information included. When information from the test is used in the feedback message, such as representing the question along with the correct response, it is generally referred to as task-specific feedback. The feedback message can also include a re-presentation of the original study materials, such as a text read before the test, in which case feedback is called instruction-based. Finally, additional information that was not presented during the initial learning or the subsequent test can be included in the feedback message, such as relating the to-be-learned material to a familiar concept from

another knowledge domain. This type of feedback is called extra-instructional. The form of feedback refers to the degree of similarity (both structural and superficial) between the original test question and the feedback message. For example, if a multiple-choice item is followed by feedback that consisted of a re-presentation of that item (including all of the potential alternatives) with the correct answer highlighted, the form of the feedback message would be essentially unchanged. However, if the order of the alternatives were re-arranged, it would constitute an elaboration of feedback form. Finally, the term feedback load is used to describe the amount of information contained in the feedback message. A simple indication of right or wrong would constitute a small load, whereas a re-presentation of the original study materials would be a large load.

Perhaps the single most important piece of information that can be included in the feedback message is the correct response. When an incorrect response is made on a test, providing the learner with the correct response allows that error to be corrected. For this reason, correct answer feedback (the re-presentation of the question stem along with the correct response) almost always results in superior learning relative to simple right/wrong feedback. For example, Pashler et al. (2005) had subjects study a list of 20 Luganda-English word pairs (e.g., leero – today) and then tested them twice on each pair (leero – ?) during an initial learning session. After each response on the initial learning tests, subjects received either no feedback, right/wrong feedback, or correct answer feedback. When subjects were tested again a week later, they produced a significantly higher proportion of correct responses in the correct answer condition relative to the right/wrong and no feedback conditions (which did not differ statistically). The advantage of correct answer feedback over right/wrong feedback replicates many other studies in the feedback literature (e.g., Gilman, 1969; Roper, 1977). In addition, the lack of a benefit of providing right/wrong feedback relative to no feedback is also a common finding (for a meta-analysis see Bangert-Drowns, Kulik, Kulik, & Morgan, 2001). Presumably, this outcome occurs because knowledge that a particular response is incorrect does not help the learner correct the error, unless there is the opportunity to re-study the material or some other way of determining the correct answer. Thus, one clear implication of the feedback literature is that the feedback message must at least include the correct response in order to promote successful learning.

Although there is consensus with regard to the importance of including the correct response, the effectiveness of further expanding the feedback message along the dimensions of type, form, and load varies considerably. Some ways of expanding upon the basic feedback message appear to have a positive effect on learning, whereas others have no effect or, occasionally, a negative effect. We provide a few examples of expansions that have yielded positive effects. Along the type of feedback dimension, some studies have found that a self-grading feedback procedure in which students are

re-presented with the original materials or required to correct their own test is superior to simply providing the correct answer (e.g., Huelser, Fazio, & Marsh, under review; Andre & Thieman, 1988). Presumably, this effect results from the active processing of the material engendered in the self-grading task, which may boost the retention of information relative to the passive processing involved in receiving correct answer feedback. In terms of the form, feedback messages that require the subject to generate the correct answer by rearranging letters have been found to be more effective than re-presenting the question and correct answer in the original form (Lhyle & Kulhavy, 1987). For the feedback load dimension, the best example may be the benefit of correct answer feedback relative to right/wrong feedback described above. Expansion of the feedback message beyond the inclusion of the correct response has yielded mixed results. Some studies have found that including an elaborative explanation of why a particular response is wrong improves learning (Peeck, 1979; Phye, 1979), but others have found no effect of elaboration (e.g., McDaniel & Fisher, 1991; Kulhavy et al., 1985). It may be that the benefits of increasing the feedback message load depend on the nature of the to-be-learned material. Correct answer feedback may be all that is needed for relatively discrete, factual material (e.g., names, dates, foreign language vocabulary, etc.), whereas learning more complex materials (e.g., concepts, applications, etc.) may necessitate an elaborate explanation of the correct answer.

The primary function of feedback is to correct memory errors (Bangert-Drowns, Kulik, Kulik, & Morgan, 2001; Kulhavy & Stock, 1989). When people retrieve information from memory, they can make two kinds of errors: (1) errors of commission and (2) errors of omission. Errors of commission occur when incorrect information is retrieved and volunteered under the assumption that it is correct (or as a guess). Errors of omission occur when information that should be reported is either retrieved but not volunteered because it is presumed to be incorrect (or irrelevant) or not retrieved at all because it is inaccessible in memory. Feedback plays a critical role in correcting both these types of errors. After an error of commission, feedback allows the learner to replace the incorrect response with the correct response. When an error of omission is made, feedback informs the learner of the information that should have been reported and allows that information to be encoded if it was not previously known.

With the ability to correct both these types of errors, feedback is a powerful tool in promoting the retention of information after tests. For example, Butler and Roediger (in press) had subjects study prose passages on a variety of topics and then take a multiple-choice test on information from the passages. Correct answer feedback was provided immediately after some questions, while no feedback was given for other questions. One week later, subjects took a final short answer test that included questions previously tested on the initial multiple-choice test as well as new questions

about information from the passages that were not tested earlier to serve as a control condition. The results showed that taking the initial multiple-choice test without feedback led to a higher proportion of correct responses on the final test relative to the no test condition, another example of the testing effect. However, taking an initial test without feedback also led to a higher proportion of intrusions (incorrect responses that were used as lures on the initial multiple-choice test) on the final test than in the no test condition. Nevertheless, the test with feedback condition led to best performance, both increasing the proportion of correct response and decreasing the proportion of intrusions on the final test.

Another function of feedback is to correct metacognitive errors. Metacognition refers to what people know about their own cognitive processes. When people retrieve information from memory, they assess the correctness of that information, a process referred to as metacognitive monitoring (Barnes, Nelson, Dunlosky, Mazzone, & Narens, 1999; Koriat & Goldsmith, 1996; Nelson & Narens, 1990). The level of confidence placed in a response is a major determinant of how subsequent feedback will be utilized. For example, when feedback study time is self-paced, subjects spend more time studying high confidence incorrect responses than low confidence incorrect responses (see Kulhavy, 1977; Kulhavy & Stock, 1989), probably because they are surprised that something they were sure is right turned out to be wrong. In general, when people are asked to assess their confidence in the correctness of a response, they are generally quite accurate. This is especially true when people have the option to volunteer or withhold responses (i.e. free report) because they tend to volunteer high confidence responses and withhold low confidence responses (see Barnes et al., 1999; Kelley & Sahakyan, 2003; Koriat & Goldsmith, 1996). However, sometimes there is a large discrepancy between actual and perceived correctness of a response, such as when an incorrect response is given in high confidence or a correct response is given in low confidence. Feedback can help to correct such metacognitive errors by reducing the discrepancy between actual and perceived correctness in a response. Research has shown that high confidence incorrect responses are particularly likely to be corrected with feedback, a finding called the hypercorrection effect (Butterfield & Metcalfe, 2001; 2006). Presumably, the surprise associated with finding out a high confidence response is incorrect increases processing of the feedback and makes the correct response more memorable. Along the same lines, feedback seems to be critical in correcting the metacognitive error inherent in low confidence correct responses (Butler, Karpicke, & Roediger, under review).

In summary, laboratory research clearly shows that feedback enhances the benefits of testing. Throughout the feedback literature, the effect of providing feedback is overwhelmingly positive. The few studies that show a negative effect have generally provided feedback in a manner that undermines learning. For example, many early programmed instruction experiments, operating on the principle that

learning should be errorless, provided the both the question and feedback simultaneously (i.e. instead of waiting until after the question had been answered to provide feedback). As a result, subjects could answer the question by looking at the correct answer at the bottom of the screen rather than trying to retrieve it from memory (see Bangert-Drowns, Kulik, Kulik, & Morgan, 2001; Kulhavy, 1977). By helping to correct errors, both in memory and metacognitive monitoring, feedback enables learners to progress towards the desired state of knowledge. We now turn to describing how feedback might shape collective memories outside the laboratory.

Feedback Shapes Collective Memories

Much as the mechanisms of retrieval and spacing are critical to the creation and maintenance of collective memories, feedback may serve as a mechanism to shape collective memories into stable, schematized narratives. When individuals in a group retrieve memories for a shared event, feedback is generated by both external (e.g., other members of the group) and internal sources (i.e. each individual listening to others and checking his or her own recollections). The feedback consists of information about both the accuracy of the retrieved memories as well as important aspects of the event (details, sub-events, etc.) that may not be recalled (errors of commission and omission, respectively). Through the process of repeatedly retrieving memories and receiving feedback from the group, the critical aspects of the event are selected by the group and eventually coalesce into a collective memory.

To illustrate this process, imagine a married couple reminiscing about their wedding. On the way to the honeymoon, they begin to share their memories of the wedding day. The husband remembers how his wife looked as she walked down the aisle, the toast his father gave during the reception, and saying thank you to all the guests at the end of the night. In turn, the wife remembers their first kiss after the ceremony, the taste of the wedding cake, and dancing with her friends to a favorite song. As they go back and forth sharing memories, each one corrects any inaccuracies in the other's recall of the events (errors of commission). For example, the husband may have attributed a funny one-liner from a toast to his father, but the wife reminds him that it was his uncle that made everyone laugh with the joke. They also remind each other of certain details and sub-events that they might not have otherwise recalled (errors of omission). For example, the wife recalled walking down the isle, but the husband reminds her that she stumbled briefly because of her long wedding dress. Over the course of their marriage this process is repeated countless times: retelling the story of their wedding to a new friend, looking through their photo album, reminiscing during an anniversary dinner, etc. As a result, the particular details and sub-events that are remembered become more consistent and uniform over time until there exists a collective memory for the event that is shared by the husband and wife.

Of course, this example of feedback shaping a collective memory in a dyad represents the simplest group possible. The idea could be expanded to larger groups, such as members of a championship football team reminiscing about the big game, war veterans from the same platoon sharing memories from their military service, and residents of a town recalling a horrible snowstorm. The critical point is that during the process of repeated retrieval of an event by members of a group, individuals receive feedback about the accuracy of their memories and any important aspects of the event that they failed to recall. Presumably, much of this feedback is given directly by other members of the group during collaborative remembering, as in the example above. However, feedback can certainly be given indirectly as well. Art, books, films, television, and other manifestations of popular culture are all prime examples of such indirect feedback. For instance, most (if not all) Americans have a collective memory for the series of events surrounding the terrorist attacks in New York City and Washington, D.C. on September 11th, 2001. In this context, reading the 9/11 Commission Report or watching the controversial film *Fahrenheit 9/11* might be an indirect way for some members of the group to influence the way in which other members of the group remember the events of 9/11. As in laboratory research, the feedback that helps to shape collective memories can take on many forms and be delivered by numerous different methods, but its function is still to correct both errors of commission and omission.

To date, no research has investigated the influence of feedback on collective memories outside the laboratory. Hence, our extensions of research to this arena as in the previous paragraphs are speculative. However, some research exists on remembering in groups (usually two or three people) that suggests that collaborative remembering facilitates subsequent individual remembering (e.g., Basden, Basden, & Henry, 2000; Weldon & Bellinger, 1997). For example, Rajaram and Pereira-Pasarin (2007) had subjects study pictures and words, and then gave them a recognition test in one of two conditions. In the collaborative condition, the group of three subjects would first discuss each item (i.e. indicating to each other whether or not they had studied it previously), then made an individual decision (yes / no) about the word. In the non-collaborative condition, subjects just made a decision about the item without any discussion. The results indicated that when subjects had the opportunity to discuss the items in a group prior to making their decision, they performed better on the recognition test than when tested individually. Presumably, subjects in the collaborative group benefited from the feedback that was generated during the discussion of each item by allowing them to correct errors of commission and omission.

Accuracy and Inaccuracy in Collective Memories

The accuracy of collective memories is a major issue for historians, anthropologists, psychologists, and the many other groups of researchers who work on this topic – not to mention the groups of people who share these memories. At best, agreement about the accuracy of a collective memory can unite a group, such as the yearly commemoration of the Declaration of Independence by Americans on the 4th of July. At worst, disputes about the accuracy of a collective memory between groups can have dire consequences. For example, a centuries-old disagreement between the Sunni and Shiite Muslims about the legitimate successors of the Prophet Mohammed continues to spark violence. Whereas Shiite Muslims maintain that the Prophet Mohammed explicitly designated his son-in-law, Ali, as successor, the Sunnis claim that Mohammed wished that his successor be appointed by the Muslim community. Regardless of whether there is consensus or disagreement about the accuracy of a collective memory, it is quite difficult (if not impossible) to determine the truth in any objective sense. In addition, even if consensus exists, the agreed upon version of events can change over time, like in the example of history textbooks given above. Thus, we will put aside the question of whether collective memories are generally accurate or inaccurate, and instead briefly discuss how the mechanisms of retrieval, spacing, and feedback described in this chapter might influence the accuracy collective memories.

Retrieval of information from memory helps people to better remember that information over long periods of time. However, taken alone, this mechanism for promoting long-term retention is completely indifferent to the accuracy of information. If incorrect information is retrieved, then memory for that incorrect information is strengthened. For example, Roediger, Jacoby, and McDermott (1996) had subjects view slides depicting a crime and then read a narrative from another witness that contained misinformation. Immediately after the narrative, they were given a forced recall test and told to recall information either only from the slides or from both the slides and the narrative. Subjects who were asked to recall information from either source produced more misinformation than those in the slides-only condition on this immediate test. A final recall test was given two days later with instructions to recall information from the slides only. Across both conditions, misinformation from the narrative that was (falsely) recalled on the initial test was more likely to be recalled on final test. That is, the act of producing information on the forced recall test seemed to strengthen the memories (the testing effect) and led subjects to be more likely to erroneously recall the information later even when they were told to be sure only to recall information from the slides (see too McDermott, 2006).

Another example comes from the literature on multiple-choice testing: when subjects (incorrectly) select a lure item, they often will produce that lure again on a subsequent test (e.g., Butler et al., 2006; Roediger & Marsh, 2005) unless feedback is given (Butler & Roediger, in press). Critically, this effect is not due to simple ignorance: the lure is more likely to be retained and reproduced later if it is selected through faulty reasoning than as a random guess (Huelser & Marsh, 2006). Overall, these findings clearly show that retrieving incorrect information on a test can lead to the learning and retention of that incorrect information. Thus, it is the accuracy of the to-be-retrieved information that determines the accuracy of the collective memories that are created.

Just as the mechanism of retrieval is blind to accuracy of information, repeated retrieval (especially when spaced) serves to increase retention of information, regardless of its accuracy. If incorrect information is repeatedly retrieved over time, it will lead to an even stronger memory for that incorrect information. Although no study has specifically investigated the effects of repeated retrieval of incorrect information, many studies have looked at how repeated exposure to misinformation influences subsequent memory. For instance, using a paradigm similar to Roediger et al (1996), Zaragoza and Mitchell (1996) varied the number of exposures to misinformation after viewing slides of a crime. Subjects were presented with the misinformation either once or three times. A greater number of presentations of the misinformation led to the production of more misinformation on a final recall test. In fact, the effect of repeated exposure to misinformation is so strong that subjects often retain the misinformation even if it is subsequently corrected (Lewandowsky, Stritzke, Oberauer, & Morales, 2005) or explicitly labeled at the time of exposure (Skurnik, Yoon, Park, & Schwartz, 2005). One explanation for this effect is that repeated exposure to misinformation increases the fluency with which that misinformation can be retrieved from memory. After a long delay, people forget that the information has been labeled incorrect, but the fluency remains and people attribute this fluency to the correctness of the information. These findings suggest that the repeated retrieval involved in creating collective memories will increase the retention of incorrect information in the exact same way as correct information. Again, the accuracy of the repeatedly retrieved information is the critical component, not the mechanism of spacing.

In contrast to retrieval and spacing, feedback can be a mechanism for reducing inaccuracy. As described above, feedback enables people to correct memory errors as they engage in repeated retrieval and form collective memories. However, feedback will only help to correct errors if the feedback succeeds in labeling the errors as incorrect. Moreover, if the information from the source of the feedback is incorrect, feedback can actually cause people to acquire incorrect information. Roediger, Meade, and Bergman (2001) conducted a study that illustrates this point. Using a relatively

new paradigm for studying false memories, they had pairs of people study six common household scenes (e.g., a kitchen). One person was a naive subject and the other was a confederate of the experimenter with instructions on how to respond. After studying a scene, each person took turns recalling items from the scenes. During this collaborative recall phase, the confederate reported either true items only or both some true items and some erroneous items, ones that did not appear any of the scenes. Later, when subjects were again asked to recall items from the scenes, they reported more erroneous items if the confederate had suggested erroneous items during the collaborative recall phase, an effect they termed the “social contagion of memory.” Subjects in this study treated recall by the confederate as feedback about items that they had not recalled. Unaware that some of the items recalled were erroneous, they used this feedback to correct what they perceived to be errors of omission. Accordingly, feedback can promote accuracy in collective memories, but only when information can be identified as incorrect. If incorrect information is deemed to be correct by other members of the group, feedback will help to promote the retention of that incorrect information and, as a result, inaccuracy in collective memories. Some collective memories may have errors introduced in this way, when one vocal member of the group remembers an event and gets some details or facts wrong.

In conclusion, the mechanisms described above each serve to promote the creation, retention, and composition of collective memories, irrespective of their accuracy. As previously discussed, one source of information that contributes to the creation of collective memories is history textbooks used in schools. If these textbooks contain inaccuracies, either intentionally or unintentionally, then students who must learn the information in those textbooks will repeatedly retrieve incorrect information and receive feedback from teachers (among other sources) that suggest that the information is correct. The result will be an inaccurate collective memory, because the mechanisms of retrieval, spacing, and feedback are largely blind to the accuracy of the information placed into the system. Although many of the inaccuracies included in history textbooks are minor or simply a matter of opinion, some inaccuracies are so blatant that they might constitute misinformation. For instance, the Kremlin has recently begun to promote a textbook entitled *A Modern History of Russia: 1945-2006: A Manual for History Teachers* that presents a new view of recent Russian history, which, among other revisions, justifies Stalin’s dictatorship as a “necessary evil.” Another example is the genocide of Native Americans by European settlers, which was largely omitted from (or else finessed) in American history textbooks until recent years. Thus, the accuracy of the information that is repeatedly retrieved and confirmed by feedback is the prime determinant of the accuracy of the subsequent collective memory.

Conclusion

In some ways, the study of collective memory is still in its beginning phases. Much has been written about the topic, but systematic exploration of the issues and the relation of issues in collective memory to the empirical study of memory by psychologists are just starting. Our aim in this chapter has been to show how lessons learned from laboratory experiments may be extended to help our framing of issues in the field of collective memory. We have discussed how repeated retrieval and feedback, in particular, are critical topics in both laboratory and real-world settings. Repeated retrieval and the feedback one receives from others in discussing some event are critical in developing what Wertsch (2002) calls schematic narrative templates that people use to organize and retrieve their collective memories.

References

- Andre, T., & Thieman, A. (1988). Level of adjunct question, type of feedback, and learning concepts by reading. *Contemporary Educational Psychology, 13*, 296-307.
- Balota, D. A., Duchek, J. M., & Logan, J. M. (2007). Is expanded retrieval practice a superior form of spaced retrieval? A critical review of the extant literature. In J. S. Nairne (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roediger, III*. New York: Psychology Press.
- Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. C. (1991). Effects of frequent classroom testing. *Journal of Educational Research, 85*, 89-99.
- Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*, 213-238.
- Barnes, A. E., Nelson, T. O., Dunlosky, J., Mazzone, G., & Narens, L. (1999). An integrative system of metamemory components involved in retrieval. In D. Gopher & A. Koriat (Eds.), *Attention and Performance XVII – Cognitive regulation of performance: Interaction of theory and application* (pp. 287-313). Cambridge, MA: MIT Press.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.
- Basden, B. H., Basden, D. R., & Henry, S. (2000). Costs and benefits of collaborative remembering. *Applied Cognitive Psychology, 14*, 497-507.
- Berntsen, D., & Thomsen, D. K. (2005). Personal Memories for Remote Historical Events: Accuracy and Clarity of Flashbulb Memories Related to World War II. *Journal of Experimental Psychology: General, 134*, 242-257.
- Bjork, R. A. (1975). Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In R. L. Solso (Ed.), *Information processing and cognition*. New York: Wiley.
-
- Brokaw, T. (2001). *The greatest generation*. New York: Random House.
- Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5*, 73-99.
- Butler, A. C., Karpicke, J. D., & Roediger, H. L., III (under review). Correcting a meta-cognitive error: Feedback enhances retention of low confidence correct responses.
- Butler, A. C., Marsh, E. J., Goode, M. K., & Roediger, H. L., III (2006). When additional multiple-choice lures aid versus hinder later memory. *Applied Cognitive Psychology, 20*, 941-956.
- Butler, A. C., & Roediger, H. L., III (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology, 19*, 514-527.
- Butler, A. C., & Roediger, H. L., III (in press). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & Cognition*.
- Butler, D. L., & Winne, P. H. (2005). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*, 245-281.
- Butterfield, B., & Metcalfe, J. (2001). Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1491-1494.
- Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition & Learning, 1*, 69-84.
- Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition, 20*, 633-642.
- 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*, 354-380.
- Chan, C. K., McDermott, K. B., & Roediger, H. L., III (2006). Retrieval induced facilitation: Initially nontested material can benefit from prior testing. *Journal of Experimental Psychology: General, 135*, 533-571.
- Conway, M. A. (1995). *Flashbulb memories*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Cull, W. L. (2000). Untangling the benefits of multiple study opportunities and repeated testing for cued recall. *Applied Cognitive Psychology, 14*, 215-235.
- Cull, W. L., Shaughnessy, J. J., & Zechmeister, E. B. (1996). Expanding understanding of the expanding-pattern-of-retrieval mnemonic: Toward confidence in applicability. *Journal of Experimental Psychology: Applied, 2*, 365-378.

- Curci, A., & Luminet, O. (2006). Follow-up of a cross-national comparison on flashbulb and event memory for the September 11th attacks. *Memory, 14*, 329-344.
- Dempster, F. N. (1989). Spacing effects and their implications for theory and practice. *Educational Psychology Review, 1*, 309-330.
- Dudai, Y. (2004). *Memory from A to Z: Keywords, concepts and beyond*. Oxford, UK: Oxford University Press.
- Ebbinghaus, H. (1967) *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Dover. (Original work published 1885).
- Er, N. (2003). A new flashbulb memory model applied to the Marmara earthquake. *Applied Cognitive Psychology, 17*, 503-517.
- Foos, P. W., & Fisher, R. P. (1988). Using tests as learning opportunities. *Journal of Educational Psychology, 80*, 179-183.
- Gates, A. I. (1917). Recitation as a factor in memorizing. *Archives of Psychology, 6*(40).
- Gilman, D. A. (1969). Comparison of several feedback methods for correcting errors by computer-assisted instruction. *Journal of Educational Psychology, 60*, 503-508.
- 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*, 392-399.
- Halbwachs, M. (1980). *The collective memory*. New York: Harper & Row. (translated by Francis J. Didier, Jr. and Vida Yazdi Ditter).
- Halbwachs, M. (1992). *On collective memory*. Chicago: University of Chicago Press. (edited, translated, and with an introduction by Lewis A. Coser).
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*, 81-112.
- Huelser, B. J., Fazio, L. K., & Marsh, E. J. (under review). Receiving feedback versus reviewing materials: Consequences for learning.
- Huelser, B. J., & Marsh, E. J. (2006, November). Does guessing on a multiple-choice test affect later cued recall? Poster presented at the Annual Meeting of the Psychonomic Society, Houston, TX.
- Jones, H. E. (1923-1924). The effects of examination on the performance of learning. *Archives of Psychology, 10*, 1-70.
- Kang, S. H. K., McDermott, K. B. & Roediger, H. L., III (2007). Test format and corrective feedback modulate the effect of testing on memory retention. *The European Journal of Cognitive Psychology, 19*, 528-558.
- Karpicke, J. D., & Roediger, H. L., III (2007a). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning, Memory and Cognition, 33*, 704-719.
- Karpicke, J. D., & Roediger, H. L., III (2007b). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language, 57*, 151-162.
- Kelley, C. M., & Sahakyan, L. (2003). Memory, monitoring and control in the attainment of memory accuracy. *Journal of Memory and Language, 48*, 704-721.
- Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in strategic regulation of memory accuracy. *Psychological Review, 103*, 490-517.
- Kramer, J. (1996). *The politics of memory: Looking for Germany in the New Germany*. New York: Random House.
- Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research, 47*, 211-232.
- Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review, 1*, 279-308.
- Kulhavy, R. W., White, M. T., Topp, B. W., Chan, A. L., & Adams, J. (1985). Feedback complexity and corrective efficiency. *Contemporary Educational Psychology, 10*, 285-291.
- Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625-632). London: Academic Press.
- Lewandowsky, S., Stritzke, W., Oberauer, K., & Morales, M. (2005). Memory for fact, fiction, and misinformation: The Iraq War 2003. *Psychological Science, 16*, 190-195.
- Lhyle, K. G., & Kulhavy, R. W. (1987). Feedback processing and error correction. *Journal of Educational Psychology, 79*, 320-322.

- Lindaman, D., & Ward, K. (2004). *History lessons: How textbooks from around the world portray U.S. history*. New York: The New Press.
- Liu, J. H., Goldstein-Hawes, R., Hilton, D., Huang, L-L., Gastardo-Conaco, C., Dresler-Hawke, E., Pittolo, F., Hong, Y-Y., Ward, C., Abraham, S., Kashima, Y., Kashima, E., Ohashi, M. M., Yuki, M., & Hidaka, Y. (2005). Social representations of events and people in world history across 12 cultures. *Journal of Cross-Cultural Psychology, 36*, 1-21.
- Logan, J. M., & Balota, D. A. (in press). Expanded vs. equal interval spaced retrieval practice: Exploration of schedule of spacing and retention interval in younger and older adults. *Aging, Neuropsychology, and Cognition*.
- Loewen, J. (1995). *Lies my teacher told me: Everything your American history textbook got wrong*. New York: W.W. Norton and Company.
- Luminet, O., Curci, A., Marsh, E. J., Wessel, I., Constantin, T., Gencoz, F., & Yogo, M. (2004). The cognitive, emotional, and social impacts of the September 11 attacks: Group difference in memory for the reception context and the determinants of flashbulb memory. *The Journal of General Psychology, 131*, 197-224.
- McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology, 19*, 494-513.
- McDaniel, M. A., & Fisher, R. P. (1991). Tests and test feedback as learning sources. *Contemporary Educational Psychology, 16*, 192-201.
- McDaniel, M. A., Roediger, H. L., III, & McDermott, K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review, 14*, 200-206.
- McDermott, K. B. (2006). Paradoxical effects of testing: Repeated retrieval attempts enhance the likelihood of later accurate and false recall. *Memory & Cognition, 34*, 261-267.
- 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.
- Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about *Challenger*. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of flashbulb memories* (pp. 9-31). 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.
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125-141). New York: Academic Press.
- Novick, P. (1999). *The Holocaust in American life*. Boston: Houghton Mifflin Company.
- Pashler, H., Cepeda, N. J., Wixted, J. T., & Rohrer, D. (2005). When does feedback facilitate learning of words? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 3-8.
- Peeck, J. (1969). Effects of differential feedback on the answering of two types of questions by fifth- and sixth-graders. *British Journal of Educational Psychology, 49*, 87-92.
- Phye, G. D. (1979). The processing of informative feedback about multiple-choice test performance. *Contemporary Educational Psychology, 4*, 381-394.
- Rajaram, S., & Pereira-Pasarin, L. (2007). Collaboration can improve individual recognition memory: Evidence from immediate and delayed tests. *Psychonomic Bulletin & Review, 14*, 95-100.
- Roediger, H. L., III (1985). Remembering Ebbinghaus. *Contemporary Psychology, 30*, 519-523.
- Roediger, H. L., III, Bergman, E. T., & Meade, M. L. (2000). Repeated reproduction from memory. In A. Saito (Ed.), *Bartlett, culture, and cognition*. Cambridge, UK: Psychology Press.
- Roediger, H. L., III, Dudai, Y., & Fitzpatrick, S. M. (Eds.). (2007). *Science of memory: Concepts*. Oxford, UK: Oxford University Press.
- Roediger, H. L., III, Jacoby, D., & McDermott, K. B. (1996). Misinformation effects in recall: Creating false memories through repeated retrieval. *Journal of Memory and Language, 35*, 300-318.
- Roediger, H. L., III, & 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., III, & Karpicke, J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249-255.
- Roediger, H. L., III, & Marsh, E. J. (2005). The positive and negative consequences of multiple-choice testing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *31*, 1155-1159.
- Roediger, H. L., III, Meade, M. L., & Bergman, E. T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review*, *8*, 365-371.
- Roper, W. J. (1977). Feedback in computer assisted instruction. *Programmed Learning and Educational Technology*, *14*, 43-79
- Rubin, D. C., & Kozin, M. (1984). Vivid memories. *Cognition*, *16*, 81-95.
- Schudson, M. (1992). *Watergate in American memory: How we remember, forget, and reconstruct the past*. New York: Basic Books.
- Shapiro (2006). Remembering September 11th: The role of retention interval and rehearsal on flashbulb and event memory. *Memory*, *14*, 129-147.
- Skurnik, I., Yoon, C., Park, D. C., & Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research*, *31*, 713-724.
- Spitzer, H. F. (1939). Studies in retention. *Journal of Educational Psychology*, *30*, 641-656.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization and memory*. (pp. 381-403). New York: Academic Press.
- Ward, K. (2006). *History in the making: An absorbing look at how American history has changed in the telling over the last 200 years*. New York: The New Press.
- Wegner, D. M. (1986). Transactive memory: A contemporary analysis of the group mind. In M. B. & G. G. R. (Eds.), *Theories of Group Behavior* (pp. 185-205). New York: Springer-Verlag.
- Weldon, M. S., & Bellinger, K. D. (1997). Collective memory: Collaborative and individual processes in remembering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1160-1175.
- Wertsch, J. W. (2002). *Voices of collective remembering*. Cambridge, UK: Cambridge University Press.
- Wertsch, J. W. (in press). Collective memory. In J. H. Byrne (Ed.), *Learning and Memory: A Comprehensive Reference*. Oxford, UK: Elsevier Ltd.
- Wertsch, J. V., & Roediger, H. L., III (in press). Collective memory: Conceptual foundations and theoretical approaches. *Memory*.
- Wheeler, M. A., & Roediger, H. L., III. (1992). Disparate effects of repeated testing: Reconciling Ballard's (1913) and Bartlett's (1932) results. *Psychological Science*, *3*, 240-245.
- Zaragoza, M.S., & Mitchell, K.J. (1996). Repeated exposure to suggestion and the creation of false memories. *Psychological Science*, *7*, 294-300.