



Published in final edited form as:

J Exp Psychol Gen. 2009 May ; 138(2): 161–176. doi:10.1037/a0015527.

Long-term memory for the terrorist attack of September 11: Flashbulb memories, event memories, and the factors that influence their retention

William Hirst

New School for Social Research

Elizabeth A. Phelps

New York University

Randy L. Buckner

Harvard University

Andrew E. Budson

Boston University School of Medicine

Alexandru Cuc

Nova Southeastern University

John D. E. Gabrieli

Massachusetts Institute of Technology

Marcia K. Johnson

Yale University

Keith B. Lyle

University of Louisville

Cindy Lustig

University of Michigan

Mara Mather

University of Southern California

Robert Meksin

New School for Social Research

Karen J. Mitchell

Yale University

Kevin N. Ochsner

Columbia University

Daniel L. Schacter

Harvard University

Correspondence can be addressed to: William Hirst, Department of Psychology, New School for Social Research, 80 Fifth Avenue, New York, New York 10011, hirst@newschool.edu. Elizabeth A. Phelps, Department of Psychology, New York University, 6 Washington Place, New York, New York 10003, liz.phelps@nyu.edu.

Publisher's Disclaimer: The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at www.apa.org/pubs/journals/xge

Jon S. Simons
Cambridge University

Chandan J. Vaidya
Georgetown University

Abstract

More than 3,000 individuals from seven US cities reported on their memories of learning of the terrorist attacks of September 11, as well as details about the attack, one week, 11 months, and/or 35 months after the assault. Some studies of flashbulb memories examining long-term retention show slowing in the rate of forgetting after a year, whereas others demonstrate accelerated forgetting. The present paper indicates that (1) the rate of forgetting for flashbulb memories and event memory (memory for details about the event itself) slows after a year, (2) the strong emotional reactions elicited by flashbulb events are remembered poorly, worse than non-emotional features such as where and from whom one learned of the attack, and (3) the content of flashbulb and event memories stabilizes after a year. The results are discussed in terms of community memory practices.

Keywords

flashbulb memories; long-term retention; memory practices; event memory; emotional memory

Brown and Kulik (1977) suggested the term *flashbulb memory* for the “circumstances in which one first learned of a very surprising and consequential (or emotionally arousing) event,” for example, hearing the news that President John Kennedy had been shot. Since Brown and Kulik’s description of their findings, the range of topics addressed in studies of flashbulb memories has grown substantially, from initial questions about special mechanisms (McCloskey, Wible, & Cohen, 1988; Neisser & Harsh, 1992) to more recent questions about the impact of aging and dementia (Budson, Simons, Sullivan, Beier, Solomon, Scinto, et al., 2004; Budson, Simons, Waring, Sullivan, Hussoin, & Schacter, 2007; Davidson, Cook, & Glisky, 2005), the history of post-traumatic stress disorder (Qin, Mitchell, Johnson, Krystal, Southwick et al., 2003), as well as the role of social identity [e.g., as seen in the presence or absence, respectively, of flashbulb memories of French citizens and French-speaking Belgians of the death of French President Mitterrand (Curci, Luminet, Finkenauer, & Gisle, 2001; see also Berntsen, 2008; Hirst & Meksin, 2008)]. Researchers have also begun to investigate memories for the flashbulb event itself (Curci & Luminet, 2006; Luminet, Curci, Marsh, Wessel, Constantin, Genocoz, et al., 2004; Pezdak, 2003; Shapiro, 2006; Tekcan, Berium, Gülgöz, & Er, 2003). In this literature, the term *flashbulb memory* refers to memory for circumstances in which one learned of the event and would include memories of where, when, and from whom one learned of, for instance, the terrorist attack of September 11, 2001. The term *event memory* refers to memory for facts about the flashbulb event and would include, for instance, that four planes were involved in the 9/11 terrorist attack and that both the Pentagon and the World Trade Center were targets.

¹There is much terminological confusion in the literature. First, the term *flashbulb memory* could be construed as implying an accurate representation of the circumstances in which one learned of the emotionally charged public event. Although we use this term here, we do not mean to imply that the memories are accurate. Second, as used in the flashbulb memory literature, the phrase *event memories* is rarely qualified, but it is not meant to refer to all event memories, only those that involve events that elicit flashbulb memories. One should more accurately refer to *memories for flashbulb events*. However, the phrase can, with repetition, become awkward and hence we adopt the convention of referring to *memories for flashbulb events* simply as *event memories*.

Flashbulb memories and their associated event memories are often considered special because they involve events that are not ordinary or everyday, and usually are not personally experienced, but rather, they are public and emotionally charged (Neisser, 1982). It is the public nature of flashbulb memories and their associated event memories that ensures the memories strongly influence both individual and collective identity (Berntsen, 2008; Hirst & Meksin, 2008; Neisser, 1982). Their role in shaping identity depends, of course, on their being retained (Bruner, 1990; Conway & Pleydell-Pearce, 2000). Surprisingly, whereas much is known about how well flashbulb and event memories are retained over a period of approximately a year, much less is known about their long-term retention. This relative neglect applies not just to the issue of the amount retained, but also to differences in the kind of information that is retained over the long-term and the factors that might affect the level and content of long-term retention. For instance, whereas many researchers have emphasized that flashbulb events inevitably elicit strong emotions from individuals, few researchers have contrasted the long-term retention of memories of these emotional reactions with the long-term retention of memories of other features of flashbulb memories, for example, who you were with when learning of the event, where you were, or how you were informed (see, however, Levine, Safer, & Lench, 2006). Moreover, although a number of psychological studies have related the level of retention to individual cognitive factors (e.g., rehearsal), none have discussed the contribution of memory practices, that is, the way a society goes about ensuring that a public event will never be forgotten by the public (Hirst & Manier, 2008; Olick & Robbins, 1998; but see Hoskins, 2007). Memory practices may play a role in the retention of flashbulb and event memories given the public nature of the reference event.

The present paper, then, focuses on four issues: (1) the long-term retention of flashbulb and event memories, (2) the comparative retention of emotional reactions with the retention of other features of a flashbulb event, (3) possible difference in the underlying processing associated with the formation and retention of flashbulb and event memories, and (4) the factors that shape long-term retention, including the role of memory practices. It explores these issues in the context of the terrorist attack of September 11, 2001.

Consider the issue of long-term retention. From the extant research, it is not clear whether forgetting for flashbulb and event memories slows or accelerates after the first year. Three studies suggest that the rate of forgetting of flashbulb memories slows dramatically after the first year. Two of these studies based their conclusions on the vividness or accuracy of flashbulb memories. Kvavilashvili, Mirani, Schlagman, and Kornbrot (2003) found that British citizens reported vivid, confidently held memories of the circumstances under which they learned of the death of Princess Diana, even after a delay of 51 months. Berntsen and Thomsen (2005) discovered that elderly Danes accurately remembered the weather on the day of the German W.W. II invasion of and withdrawal from Denmark.

Neither of these studies, however, employed a test-retest methodology, in which memories are assessed shortly after the flashbulb event and then after a significant retention interval. This methodology supplies a putatively reliable memory with which to compare the consistency of later recollections and is consequently the preferred means of studying flashbulb memories (see, however, Winningham, Hyman, & Dinnel, 2000). Kvavilashvili et al. did not have an initial assessment for a test-retest. Berntsen and Thomsen had verifiable information about the original event, but their documentary methodology does not permit as wide a ranging examination of mnemonic attributes as the test-retest method does. Relying on public records such as weather reports largely precludes exploring those attributes Brown and Kulik (1977) identified as the canonical features of flashbulb memories, for example, who the respondent was with, how the respondent reacted emotionally, or who the informant was. Bohannon and Symons (1992; see also Bohannon, 1988) conducted the third study,

finding a slowing in forgetting, and did employ a test-retest methodology in their investigation of the Challenger explosion, but, in the end, based their conclusions about the rate of forgetting on cross-sectional data.

Two studies did ground their conclusions about long-term retention on the results of test-retests. Unfortunately, Neisser and Harsh (1992) employed only one retest in their study of the Challenger explosion, making any analysis of the rate of forgetting difficult. On the other hand, Schmolck, Buffalo, and Squire (2000) used two retests in their study of the announcement of the verdict in the O.J. Simpson criminal trial. They found that at 15 months, a little less than 40% of the flashbulb memories they examined contained no distortions, and only about 10% contained major distortions. At 32 months, the pattern was reversed: Only about 20% contained no distortions and over 40% of the memories contained major distortions. These results strongly support the claim that the rate of forgetting increases, rather than slows, over time.

Because of the controversy over the rate of forgetting of flashbulb memories, it is difficult to evaluate Talarico and Rubin's (2003) claim that, despite a flashbulb event's public and emotionally charged nature, the rate of forgetting of flashbulb memories is the same as the rate of forgetting of ordinary autobiographical memories. Talarico and Rubin suggested that flashbulb memories and ordinary autobiographical memories differ not in their rate of forgetting, but in the confidence with which they are held, with confidence in flashbulb memories remaining high, even as the memories are forgotten. Confidence in ordinary autobiographical memories declines as the memories are forgotten (see also Weaver, 1993). Talarico and Rubin, however, only tested retention intervals of eight months or less. Schmolck et al. (2002)'s findings indicated that flashbulb memories may be an exception to the pattern of forgetting observed for ordinary, autobiographical memories when long-term retention intervals are considered.

Talarico and Rubin (2003) compared their participants' memory for their reception event for 9/11 with a self-selected autobiographical memory – a memory of an “everyday” event from the three days before September 11. A perhaps more general point of comparison would be the forgetting curves obtained in diary studies (Rubin, 2005). These studies involve the assessment of a wide-range of types of memories over a substantial period. The forgetting curves collected across studies are remarkably similar, showing rapid forgetting in the first year and then slowing. As a result, they indicate that autobiographical memories may follow the well-established pattern of forgetting documented since Ebbinghaus (1913/1964). Linton (1986), for instance, showed dramatic forgetting over the first year and then a much slower rate of forgetting of 6% for the next five years. Similarly, Wagenaar (1986) found a substantial decline of 20% in the first year for critical details and then a slower decline of approximately 10% for the next four years. If Talarico and Rubin's findings of equivalent forgetting of flashbulb and ordinary memories up to eight months extends to longer retention intervals, then the diary studies would suggest that the results of Schmolck et al. (2000) are an anomaly, and those of Bohannon and Symons (1992), Kvavilashvili et al. (2003), and Berntsen and Thomsen (2005) may be more typical.

In the present study, we asked whether the accelerated forgetting Schmolck et al. (2000) observed for the Simpson verdict between the first and third years applies as well to flashbulb memories for the terrorist attack of 9/11. Consequently, we assessed our participant's memory for 9/11 one week, 11 months and 35 months after the terrorist attack. We choose the 11 months and 35 months retention intervals because they were in the same time frame used by Schmolck et al., but minimized potential effects of anniversary commemorations.

In addition, we also examined the retention of associated event memories at one week, 11 months, and 35 months after the terrorist attack. The scant relevant literature on event memory is as inconclusive about long-term retention as the literature on flashbulb memories. Bahrick and his colleagues (Bahrick, 1983, 1984; Bahrick, Bahrick, & Wittlinger, 1975) have shown that neutral facts, such as the names of fellow college students, college streets, or college-learned Spanish vocabulary, are steadily forgotten for six years, and then, if still retained, preserved for decades to come. Along the same lines, Belli, Schuman, and Jackson (1997) found good retention after decades for newsworthy events such as the Tet Offensive, at least for participants for whom the event “defined” their generation. Neither of these studies examined whether respondents remembered the circumstances in which they learned of the event, making their relevance to the topic of flashbulb memories at best speculative. The flashbulb memory studies that also explored event memory indicate that for retention intervals of a year or less, event memories are subject to a steady decline (Finkenauer, Luminet, Gisle, & Filipopot, 1998; Smith, Bibi, & Sheard, 2003; Tekcan et al., 2003). In the only study that examined event memory at longer retention intervals, Bohannon and Symons (1992) found that event memories declined a substantial 20% between the 15-month and 36-month intervals, suggesting that, while the rate of forgetting may not accelerate, it clearly does not slow after a year.

Finally, as noted earlier, we also investigated whether long-term retention for emotional reactions to 9/11 differs from memory for other features of flashbulb memories, as well as what factors shape long-term retention of different features. There is almost no research comparing memory for emotional reactions with memory for other features of flashbulb memory (but see Qin et al., 2003). Levine and colleagues explored memory for emotional reactions to flashbulb events, but used relatively short retention intervals and did not make comparisons with other features, as we do here (Levine, Prohaska, Burgess, Rice, & Lahlhere, 2001; Levine, Whalen, Henker, & Jamner, 2005).

As for the factors that might affect retention, we explore whether any similarity in the patterns of forgetting of flashbulb memories and event memories implies that the processes that underlie the retention (and forgetting) of these two types of memories are the same. Some research on flashbulb memories and event memories would suggest that the underlying processes are similar. For instance, through their modeling of the formation and maintenance of flashbulb memories, Luminet, Curci, and their colleagues have argued that some factors involved in the formation of event memories overlap with those involved in the formation of flashbulb memories (e.g., rehearsal, see Luminet, 2008, for a review). Their model also documented differences in factors uniquely involved in the formation of flashbulb memories, specifically, surprise and novelty. Testing the complex models developed by this group goes beyond a chief aim of the present paper – to explore long-term retention of flashbulb memories and event memories. Nevertheless, we investigated some factors that could putatively predict levels of forgetting. We also examined the way the content of the memories changes over time, on the assumption that if the predictors or content changes differ for flashbulb and event memories, then different processes may be involved.

We are not only interested in intrapsychic factors. We also present analyses in the General Discussion that suggest provide evidence that different retention curves reported in the literature can be attributed to different social memory practices.

Method

Participants, recruitment, and procedure

Participants were recruited in Boston and Cambridge, MA; New Haven, CT; New York, NY; Washington, DC; St. Louis, MO; Palo Alto, CA; and Santa Cruz, CA. For Survey 1 (one week after the attack), tables were set up either on the campuses of the collaborators or in surrounding neighborhoods. Lab members also asked friends and acquaintances if they would be interested. In the survey, we asked if participants were willing to be contacted in a year for a second survey. All respondents indicated their willingness. Surveys and stamped return envelopes were given to all participants.

For Survey 2, we contacted respondents to the first survey through e-mail, the postal system, or both, and asked them if they wished to participate in the second survey. We also recruited additional participants, in the same ways, for both the second survey and a third survey, to examine possible effect of prior participation. For Survey 3, we recruited all participants who responded to Survey 1 and/or Survey 2 and added another group of new participants.

For all three surveys, participants were told that they had one week to fill out the survey and return it to the experimenters. There was both a paper version and a web-based version for Surveys 2 and 3. Participants recruited through the postal system received a survey, but were told they could use the web-based version if they wished. Those who were recruited through e-mail were told that they could fill out the web-based version or receive a paper version either through e-mail (as a pdf file) or in the post. We recruited participants between September 17, 2001 and September 21, 2001 for Survey 1; August 5 and August 26, 2002 for Survey 2; and August 9 and August 20, 2004 for Survey 3. We closed the website two weeks after the last day of recruitment and stopped accepting returned postal surveys five days later.

Table 1 shows the number of participants from each of the six recruitment locations. 38% of the respondents to Survey 1 completed Survey 2, while 18% of the respondents on Survey 1 completed both Surveys 2 and 3. For Surveys 2 and 3, 27% responded through e-mail. These return rates are comparable to other surveys without a monetary incentive or a follow-up query (Baruch, 1999). We compared the responses to each question on the survey, one question at a time, and found no significant differences in the responses of the web-based and postal responses (in all cases, $p > .4$), thus we merged the data from the two formats. To make the exposition in this paper straightforward, we will confine most of our analysis to the 391 participants who filled out all three surveys.

Surveys

Separate surveys were designed for each testing period, with Survey 1 serving as the model for the other two. The surveys were approximately 17 pages and took about 45 minutes to complete. Copies of the surveys can be found at <http://911memory.nyu.edu>.

All surveys began with a general statement of the aims of the project, a consent form, and a request for an identification code that would allow the experimenters to track questionnaires across the three survey periods. Table 2 summarizes the probes on the questionnaire in Survey 1 that figured in our present analyses. Questions 1 – 6 were relevant to establishing the consistency of flashbulb memories; Questions 7 – 11, the accuracy of event memories, and Questions 12 – 23 dealt with predictors, specifically, consequentiality (as assessed by personal loss or inconvenience), the intensity of the emotional response, and rehearsal (as assessed by attention to media and conversations). We did not attempt to cover the entire range of predictors found in the literature. In some cases, such as surprise and novelty, we expected uniformly high scores, making such data insensitive as a potential predictor. In

other cases, such as prior knowledge, we were uncertain what to ask as we prepared the survey a few days after the attack (e.g., while we were constructing the survey, there were still questions about who carried out the attacks). Survey 2 followed a similar format to Survey 1, except that two versions were constructed and distributed such that participants were asked, in equal numbers, for the flashbulb memory questions either: (a) How confident are you that your recollection is accurate (different questions assessed time, source, place, etc.), or (b) How accurately do you think that you will remember two years from now? Participants responded on a 1 – 5 scale, with 5 being the highest rating. Survey 3 was similar to Survey 2, although the time frame for the forecasting questions was changed from two years to seven years. Eight demographic questions concluded the surveys, probing, among other things, for residency.

Coding

A coding manual for Survey 1 was developed after reading through 50 surveys to determine the range and nature of the responses. It was written to be a stand-alone document that would provide complete and independent guidance to a coder. Table 3 contains examples of the coding scheme. If 50 similar responses were coded as “other,” then the coding scheme was revised and this “new” option was added. The coding was then redone for this question. Such recoding was done for 14% of the questions. The coding manuals can also be found on <http://911memory.nyu.edu>.

To assess interrater reliability of the coding, at the end of the coding process for each survey, we randomly selected 10% of the surveys to be dual-coded. We then calculated for each question either kappas or Cronbach alphas (whichever was appropriate) for each question. Reliability ratings were good for both the short-answer questions and open-ended questions, in that they all exceeded .80.

Results

General Considerations

As Luminet, Curci, Marsh, Wessel, Constantin, Gencoz, et al. (2004) noted, a large sample and numerous comparisons can produce misleading significant differences. Following their guidelines, we set a significance level of .01. Moreover, we report Cohen's *d* (Cohen, 1992), for which .20 is indicative of a small effect size, .50 a medium effect size, and .80 a large effect size.

Forgetting

We began by comparing the rate of forgetting a year after the September 11 attack with the rate of forgetting after three years.

Coding considerations—We devised separate coding schemes for flashbulb memories and event memories. Our coding scheme for measuring the consistency of flashbulb memories differed from the one employed in Neisser and Harsch (1992). We developed this new procedure because we wanted not simply to determine whether responses were consistent over time, but also how they varied in content from one survey to the next. In our measure of consistency of flashbulb memories, we matched the coding for Survey 1 with the coding of the other two surveys, producing consistency measures that contrasted Survey 2 with Survey 1 (S12) or Survey 3 with Survey 1 (S13). Two responses were consistent if they were coded in the same manner, with a “1” assigned if the items were consistent and a “0” if they were inconsistent. As Table 2 indicates, we focused on six of the canonical features of Brown and Kulik (1977). The six consistency scores were averaged to form an overall measure of consistency, ranging from 0 to 1.

The Neisser-Harsch coding scheme allowed for graded scoring, whereas our scheme did not. That is, in Neisser and Harsch (1992), a “correct response” could have received a score of “2” or “1,” with “0” reserved for clearly incorrect responses. Our measure was dichotomous. If a participant had originally wrote “I was listening to the TV as I got dressed” and later remembered “I was watching TV,” Neisser and Harsch would have scored it a “1” out of a possible two. We would have scored it a “1” out of a possible one (see Table 2). Consequently, when the various scores are summed over canonical features, the relative ranking of two participants might differ according to the Neisser-Harsch and our coding schemes. In order to explore the relation between these two scoring procedures, we asked two coders to follow the Neisser-Harsch scheme for 50 participants' responses to the three surveys. The coders evidenced a high degree of interrater reliability, $\kappa = .81$. The correlations between our overall consistency scores and the consistency score based on the Neisser and Harsch scheme were significant (S12: $r = .29, p < .05$; S13: $r = .38, p < .01$). Although these significant correlations are not large, they suggest that the pattern of results we observed would also have been found if we had followed the procedure specified by Neisser and Harsch. In order to assess this claim, we redid the analyses presented below in the section on “forgetting and flashbulb memories” using the scores derived from the Neisser-Harsch coding scheme. We found the same pattern of results as the one reported in this section for the Neisser-Harsch coding (in all cases, $p < .05$).

As for our coding scheme for event memory, we compared the answers to our probes about the event itself with the correct answers, as determined by news accounts. As Table 2 indicates, we probed for five different sets of facts: (1) number of planes, (2) name of airlines, (3) location of attacks, (4) location of President Bush, and (5) order of major events. With respect to the questions about the number of planes and about where President Bush was at the time, if respondents were correct, they received a score of 1. Otherwise, they received a 0. For the question about the identity of the airlines, each correctly identified of the two involved airline carriers received a score of .5. Furthermore, for each incorrectly mentioned airline carrier, we subtracted .25 from the total score, with a maximum penalty of .5. To keep the range of scoring between 0 and 1, we changed any negative score to a zero. We scored (9) in a similar manner, but since there were three crash sites, each correct response received a score of .33 and incorrect answers were penalized at a rate of .16. For (11), we had listed six possible events for the respondent to order (see Table 2). We calculated the Spearman Rank correlation between the respondent's order and the actual order. A negative correlation was recorded as 0. The total accuracy score was the average score across the five probes. Here and elsewhere, we use the term *accuracy* when discussing event memory; *consistency* when considering flashbulb memories.

Forgetting and flashbulb memories—We were chiefly interested in determining whether the rate of forgetting increased or slowed over the long term, specifically, between Survey 1 and Survey 3. As Table 4 reveals, 11 months after the attack (when Survey 2 was administered), participants offered consistent answers about their flashbulb memories only 63% of the time, on average. The decay over the next two years (when Survey 3 was administered) was much smaller, with a proportional decline of 9%, or an average of 4.5% a year. Although the difference between the consistency between Surveys 1 and 2 and the consistency between Surveys 1 and 3 was significant, $t(390) = 5.21, p < .01$, the effect size is small ($d = .28$), providing further support that the rate of forgetting had slowed after the first year.

In order to contrast memory for emotional reactions on hearing the news about 9/11 with other features of flashbulb memories, we separately tabulated the consistency of responses across the three surveys to an open-ended question about the emotional reaction of the respondent upon hearing the news (Table 2, Question 4). For both Surveys 2 and 3, the

overall measure of consistency was significantly greater than the measure of consistency associated with the open-ended probe of emotion: for S12, $t(375) = 8.72$, $d = .56$, $p < .01$, for S13, $t(364) = 9.17$, $d = .53$, $p < .01$ (see Table 4).

Participants' relatively poor recollections of their emotional state can also be detected in their responses to the six questions that specifically asked them to rate the intensity with which they felt sadness, anger, fear, confusion, frustration, and shock (Table 2, Questions 14–20). We were not interested in participants' recollection of the specific rating score they gave: This would require them to remember both the level of their intensity and the scale they used to express this intensity across surveys. Rather we were interested in their memory of the relation among different emotions, for example, whether, after 9/11, they felt more sadness than shock. To explore these relations, we translated the emotional ratings an individual participant gave into z-scores, calculated separately for each survey and each participant. We calculated Pearson Product-Moment correlations between a participant's z-scores on Survey 1 and Survey 2 and between a participant's scores on Survey 1 and Survey 3. Table 4 contains the average correlation for these two calculations. The correlation between Survey 1 and Survey 3 was significantly less than between Survey 1 and Survey 2, $t(390) = 2.90$, $p < .05$, but the effect size was small, $d = .15$, indicating that most of participants' forgetting of their emotional responses happened between Survey 1 and Survey 2. For an individual respondent, the correlation would need to be greater than .73 to be significant at the .05 level. Only 31.3% of the respondents had a correlation greater than .73 between Surveys 1 and 2 and only 25.3% between Surveys 1 and 3.

Finally, although the flashbulb memories were not consistent across surveys, confidence ratings were high (see Table 4). As noted, this pattern of inconsistent memories accompanied by high confidence rating suggests that a trademark of flashbulb memories extends across long-term retention periods (Talarico & Rubin, 2003). The decline in confidence between Survey 2 and Survey 3 was not significant ($p > .30$).

The pattern of results we found did not arise because we repeatedly tested our participants. In order to determine whether the small decline in overall consistency we observed between Survey 2 and Survey 3 could be attributed to an effect of filling out Survey 2, we compared the overall consistency scores for the sample that filled out only Surveys 1 and 2 ($M = .63$; $SD = .20$) with overall consistency scores for the sample that filled out only Surveys 1 and 3 ($M = .55$; $SD = .23$). The difference between these two consistency scores represented a significant decline, $t(458) = 3.44$, $p < .01$, again with only a medium effect size, $d = .32$. In order to explore further the effect of multiple surveys, we also compared the overall consistency scores on Survey 3 of the three-surveys sample with (1) the overall consistency scores for participants who only filled out Surveys 1 and 3 and (2) the overall consistency scores for participants who only filled out Surveys 2 and 3 (with Survey 2 now serving as the baseline). These two comparisons were not significant ($ps > .40$). In addition, there were no significant differences between the various samples in terms of age, religion, residency, political viewpoint, gender, or race/ethnicity ($ps > .30$). In other words, the large decline we observed between Surveys 1 and 2 and the smaller decline that occurred between Surveys 2 and 3 is probably not a result of our retesting procedure.

Forgetting and event memory—Similar to the consistency measure for flashbulb memories, the overall measure of event memory accuracy showed a pattern of slowing in the rate of forgetting for facts about 9/11 between the first and third year (see Table 5). Examining the overall measure of accuracy, an ANOVA revealed a main effect for survey, $F(1, 390) = 88.5$, $p < .001$, $\eta_p^2 = .19$. The drop in accuracy from Survey 1 to Survey 2 was significant, with a decline of 13% and a medium effect size, $t(390) = 11.81$, $p < .001$, $d = .$

61. We did not find a significant decline in accuracy from Survey 2 to Survey 3, $t(390) = .77, p = .45$.

A closer examination of the responses to each probe revealed a more complicated story than suggested by overall accuracy scores. As Table 5 indicates, the pattern of forgetting depended on the information being sought. There was no significant difference between the accuracy on Surveys 1 and 2 for two probes: the crash sites and the order of the events ($ps > .50$). For the number of the planes, there was only a significant decline from Survey 1 to Survey 2, using a McNemar test, $\chi^2(1) = 14.75, p < .001$.

The names of the airline carriers showed a continuous decline across surveys and large effect sizes: Survey 1 vs. Survey 2, $t(390) = 7.72, p < .001, d = .50$; Survey 2 vs. Survey 3, $t(390) = 6.21, p < .001, d = .30$. We account for this result in the General Discussion section. As for the probes about the location of President Bush at the time of the attack, again, using a McNemar test, there was a significant decline from Survey 1 to Survey 2, $\chi^2(1) = 95.43, p < .001$, as well as a significant improvement from Survey 2 to Survey 3, $\chi^2(1) = 68.81, p < .001$. We attribute the increase in accuracy about the location of President Bush from Survey 2 to Survey 3 to what we call the *Michael Moore Effect* (also see Greenberg, 2004). Michael Moore's film *Fahrenheit 911* brought dramatic attention to President Bush's location by featuring his reading of *The Pet Goat* in a Florida elementary school. Table 5 contrasts the frequency of correct responses for the question about President Bush's location for those who reported that they did or did not see the Moore film. There was no significant difference between those who did or did not watch the Moore film on Surveys 1 and 2, but a significant difference emerged on Survey 3, $\chi^2(1) = 24.41, p < .001$. In other words, there was a marked overall improvement between Surveys 2 and 3, with those who saw the movie showing a greater improvement for Survey 3 (52%) than those who had not (32%). The improvement of those who did not see the Moore film may reflect the extensive discussion in the media the Moore film generated about the six-minute segment of President Bush's *Pet Goat* reading. The even more dramatic improvement of the Moore movie watchers may have been because of the film itself.

The only significant correlation between overall consistency and overall accuracy was between the overall consistency measure between Surveys 1 and 3 and the overall accuracy measure on Survey 3, $r = .10, p < .05$.

In sum, the present results suggest that the rate of forgetting slows between the first and third years for both flashbulb memories and event memories. This result suggests that the Schmolck et al. (2000) results may be an anomaly rather than a characteristic portrait of the pattern of forgetting of flashbulb memories over the long-term. Moreover, the lower emotion consistency than overall consistency scores suggest that people may forget their emotional reactions to hearing the news of the attack more quickly than other aspects of flashbulb memories, such as, where they were, who told them, and how they were told.

Are there different processes underlying the similar patterns of forgetting for flashbulb memories and event memories?

Even though we found similar patterns of forgetting for both flashbulb memories and event memories, this does not mean that the same factors are affecting remembering and forgetting in the two cases. We address this issue by examining whether the same factors predict consistency and accuracy and whether the pattern of types of changes in content over the long-term are the same for flashbulb memories and event memories. If differences in predictors and content changes can be found between flashbulb memories and event memories, then different processes probably underlie the retention (and forgetting) of these two memory types.

Predictors of consistency and accuracy—We focused on five putative predictors: two probes of consequentiality—residency and the combination of personal loss and inconvenience – as well as emotionality, media attention, and ensuing conversation. We also examined the location in which a participant learned of the attack as a predictor, but it had no effect on our measures of consistency, confidence, or accuracy, and is not discussed further. As to residency, we divided our sample into New Yorkers and non-New Yorkers. Respondents who resided outside the city borders were classified as non-New Yorkers. We explored whether participants who lived in downtown Manhattan (near Ground Zero) differed from other participants, inasmuch as other researchers have found differences between the downtown population and the larger population (Galea & Vlahov, 2004; Sharot, Martorella, Delgado, & Phelps, 2007). We failed to find any differences on our measures of consistency, accuracy, or confidence using this distinction ($ps > .20$).

In assessing the effect of personal loss and/or inconvenience (Table 2, Questions 12 and 13), we counted concrete answers such as damage of home, loss of business, personal injury to self, friend, or relative, cancellation of school, and/or lack of food.² We did not include psychological distress as a form of loss or inconvenience (e.g., felt anxious, lost appetite), although good arguments could be made to do so. This classification scheme should not adversely skew our results. If anything, it should decrease the likelihood of finding differences that might arise because of “personal loss or inconvenience,” inasmuch as it excludes from the “loss” sample participants who reported suffering psychological distress. An individual was said to “experience personal loss or inconvenience” if they stated one “concrete” example. According to this criterion, 40.4% of the respondents who completed all three surveys experienced personal loss or inconvenience.

In assessing emotional intensity, the surveys asked participants to rate the intensity of their emotions on a 1 – 5 scale, with 5 being most intense (see Table 2, Questions 14–20). Inasmuch as we were mainly interested in the effects of participants' initial emotional reaction, we focused our attention on the responses recorded on Survey 1, deriving two measures of overall intensity from Survey 1's six emotional probes: (1) the average of the six emotions we probed for and (2) the highest rating given to the six emotions. These two measures were significantly correlated ($r = .70, p < .001$). Both scores yielded enough variability to permit further analysis.

Questions (21) and (22) in Table 2 probed for what might be treated as effects of rehearsal, in particular, the level of media attention and the degree of ensuing conversation. In addition, on Survey 1, we asked respondents to indicate how they spent their days following the attack, assigning a percentage to a list of activities (see Table 2, Question 23). We summed the percentages assigned to activities (a), (c), (d), and (e) to obtain a measure of attention to the media. The percentage associated with (b) reflected the level of ensuing conversation. The correlations between these percentage scores and the 1–5 rating of media attention and ensuing conversation were significant, respectively, $r = .29, p < .005$, and $r = .33, p < .005$. We use the Likert-scale ratings of media attention and ensuing conversation in our analyses.

We found a clear difference in the extent to which our five putative predictors correlated with the level of consistency of flashbulb memories and with the degree of accuracy of event memories. None of the five putative predictors appeared to be related to the consistency of flashbulb memories, in either Survey 2 or Survey 3 (see Tables 6 and 7). We did find a suggestion that the consistency ratings of non-New Yorkers were actually greater than those

²To be specific, we excluded from our classification of “experiencing personal loss or inconvenience,” columns 107, 108, 113–115, 117–121, and 124–126, using the column numbers from the coding manual for Survey 1.

of New Yorkers, $t(298) = 2.00, p < .05, d = .24$, a difference also found in Pezdek (2003). Inasmuch as our difference did not achieve the .01 level of significance, we remain cautious in interpreting this counterintuitive trend. As far as emotionality is concerned, we failed to find any correlations between consistency and emotionality when separately calculating correlations for each of the six emotions probed for in the surveys.

On the other hand, four of the five factors were related to the accuracy of event memory: residency, personal loss/inconvenience, ensuing conversation, and media attention. A three-way ANOVA on the three dichotomous factors – with survey, residency, and personal loss/inconvenience as the dependent variables and overall accuracy as the independent variable – revealed main effects for residency, $F(1, 296) = 4.39, p < .05, \eta_p^2 = .15$, and survey, $F(1, 296) = 38.54, p < .001, \eta_p^2 = .12$, as well as a two-way interaction between personal loss/inconvenience and residency, $F(1, 296) = 7.22, p < .01, \eta_p^2 = .15$, and a three-way interaction between survey, residence, and personal loss/inconvenience, $F(1, 296) = 4.89, p < .05, \eta_p^2 = .02$ (see Table 6). As in Pezdek (2003), the main effect for residency indicates that New Yorkers' event memory was more accurate than that of non-New Yorkers. The three-way interaction could be attributed to the failure to find any significant differences for Survey 1 ($p > .20$). The two-way interaction between personal loss/inconvenience and residency can be traced to the significant differences between those with or without personal loss/inconvenience that emerged for the non-New Yorker sample on Survey 2, $t(193) = 3.51, p < .001, d = .66$, and on Survey 3, $t(193) = 2.65, p < .01, d = .44$, a difference that did not appear in the New Yorker sample on either survey ($ps > .20$).

The relation between media attention and ensuing conversation, on the one hand, and accuracy, on the other, is revealed through correlational analyses (see Table 7). Accuracy on Survey 1 was significantly correlated with media attention and ensuing conversation in the first two weeks, but not with media attention and ensuing conversation over the 11 months or three years that followed. Accuracy on Survey 2 was correlated with initial attention to the media and initial conversations, but also the media attention and ensuing conversation over the next 11 months. Accuracy on Survey 3 was correlated with media attention and ensuing conversation in the first few days and over the 35-month period, as well as with ensuing conversation over the first 11-months.

We explored whether media attention and ensuing conversation served as mediators of our observed relation between memory accuracy and residency on the one hand, and personal loss/inconvenience on the other. We devised a measure of the cumulative level of media attention and ensuing conversation for the period of time covered by a particular survey by calculating the average of the ratings given on that survey and any previous surveys. For instance, to calculate the level of media attention relevant to Survey 3, we averaged over the rating of media attention provided in Surveys 1, 2, and 3. (For clarity's sake, *media attention* will refer to the rating participants gave to the question about how much they attended to the media on a particular survey. *Level of media attention* will refer to the averaged measure.)

Table 8 contains the levels of media attention and ensuing conversation after two weeks, 11 months, and 35 months. New-Yorkers and non-New Yorkers and those with or without a personal loss or inconvenience attended to the media equally; that is, we found no significant main effects or interactions for level of media attention ($ps > .30$). This result suggests that media attention may not be a mediating factor for residency and personal loss/inconvenience. We did, however, find differences for the level of ensuing conversation. In a three-way ANOVA between survey, residency and personal loss/inconvenience, there was a main effect for survey, $F(1, 285) = 265.34, p < .01, \eta_p^2 = .48$, and a significant interaction between residency and personal loss/inconvenience, $F(2, 285) = 3.77, p < .05, \eta_p^2 = .04$. Non-New Yorkers with a personal loss or inconvenience talked significantly more about the

attack than non-New Yorkers without a personal loss or inconvenience, for all three time periods (2 weeks: $t(192) = 1.97, p < .05, d = .28$; 11 months: $t(183) = 2.71, p < .01, d = .40$; 35 months: $t(183) = 2.92, p < .03, d = .43$). On the other hand, there was no significant difference in the conversations of those New Yorkers with or without a personal loss or inconvenience ($p > .50$). These results nicely reflect the pattern of results we found for the effects of personal loss/inconvenience and residency on accuracy.

To determine whether the level of ensuing conversation does indeed mediate the effects of residency and personal loss/inconvenience on accuracy, we conducted three mediational analyses on the non-New Yorkers' data (Baron & Kenny, 1986). The independent variable was personal loss or inconvenience (for non-New Yorkers), the dependent variable was accuracy, and the mediational variable was the level of ensuing conversation. We focused on the non New-Yorkers because that is where the level of ensuing conversation differed as a function of personal loss/inconvenience (see Figure 1). For the initial two-week period, the level of ensuing conversation failed to predict accuracy. Personal loss/inconvenience of non-New Yorkers also did not predict accuracy. Thus, basic assumptions underlying the mediational analysis were violated. According to Baron and Kenny, partial mediation is suggested if the regression coefficient associated with the path from personal loss to event memory decreases when the mediating variable is included in the regression. By this standard, we found evidence for partial mediation for the 11-months and 35-months periods. As indicated by the Sobel (1982) test, at 11 months and 35 months, the mediator of level of ensuing conversation carried the influence of personal loss to event memory (11 months: test statistic = 2.01, $p < .05$; 35 months: test statistic = 2.04, $p < .05$). These analyses suggest that what mattered was not particularly where participants lived at the time of the attack or what personal loss/inconvenience they experienced, but how much they talked about the event.

In addition to consistency and accuracy, another important variable in studies of flashbulb memories is confidence. Although our five putative predictors did not have an effect on consistency, both media attention and ensuing conversation affected the level of confidence with which participants held their flashbulb memories (see Table 6 and especially Table 7). For Survey 2, how much respondents attended to the media and talked about the attack in the first 11 months was significantly correlated with level of confidence. As for their confidence after 3 years, what mattered was how much respondents attended to the media and talked about the attack over the three years, not just in the first few days or the first year.

Changes in the content of memories over time—The analyses of putative predictors suggest that the similar patterns in the rate of forgetting we found for flashbulb and event memories involve different underlying processes. We can further buttress this claim by examining the changes in the content of memories over time: Different pattern of changes for flashbulb and event memories would suggest that different underlying processes are involved in their retention (and forgetting). First, consider those flashbulb memories that were consistent and those event memories that were accurate. Did, for instance, the consistent flashbulb memories on Survey 2 remain consistent on Survey 3? Did the accurate event memories on Survey 1 remain accurate on Survey 2? And did the accurate event memories on Survey 2 remain accurate on Survey 3? The result in Figure 2 suggests that answers to these questions are mostly positive. In examining flashbulb memories, we separated responses concerning “objective” canonical features -- place, informant, ongoing activity at the time of the reception, and the activity immediately following the reception -- from the one feature involving emotional reaction. For objective features, a consistent response on Survey 2 led 82% of the time to a consistent response on Survey 3. The results for event memory are similar, with accurate response on Survey 1 remaining accurate on Survey 2 and accurate responses on Survey 2 remaining accurate on Survey 3. The one exception to this pattern is the memories participants reported for their initial emotional

reaction. A consistent response on Survey 2 remained consistent on Survey 3 only 44% of the time. This result is in line with our finding that the relative rated strength of different emotions changed across surveys (correlation of emotion z-scores, Table 4).

What about changes to inconsistent/inaccurate responses from one survey to the next? Inconsistent/inaccurate responses in one survey could be followed up in the next survey in at least three ways: through correction, repetition, or alteration. For instance, an inconsistent response on Survey 2 (based on the responses on Survey 1) could be:

- (1) Corrected in Survey 3. This is the proportion of inconsistent responses in Survey 2 revised in Survey 3 to be consistent with what appeared in Survey 1 (the *corrected response proportion*),
- (2) Repeated in Survey 3. This is the proportion of inconsistent responses in Survey 2 (when compared to Survey 1) that were repeated in Survey 3 (the *repeated response proportion*),
- (3) Altered in Survey 3 to something other than the response on Survey 1. This is the proportion of inconsistent responses in Survey 2 remembered differently in both Surveys 1 and 3 (the *other response proportion*).

In exploring these three options for flashbulb memories and their analogs for event memories, we did not include in our analyses those instances in which participants failed to answer a probe.

Is the frequency of one of these types of change larger than the frequencies associated with the other types? Does the pattern of frequency distribution differ for flashbulb and event memories? For objective features of flashbulb memories, participants tended to repeat their inconsistencies. For these features, there were significant differences between the proportion of *repeated responses* and the proportion of *other responses*, as well as between the proportion of *repeated responses* and *corrected responses* (both *t*-tests, $p < .001$). Such repetitions are concordant with a slowing rate of forgetting between Survey 2 and Survey 3. Moreover, the presence of repetitions suggests that a stable memory is forming after a year delay, with stories about the circumstances in which one learned of the terrorist attack remaining the same over the long-term, even if they are full of inconsistencies with the initial report. In contrast, participants were inclined to report remembering an emotion not reported in their original survey or in Survey 2 over either correcting or repeating a previous response (all *t*-tests, $p < .01$). This finding, again, is in line with our report of poor emotional memory.

The changes observed for event memories differed from what we observed for flashbulb memories, with event memories corrected rather than repeated from one survey to the next. We can determine the change in the content of event memories from Surveys 1 to 2 (indicated as S2 on Figure 2) and from Surveys 2 to 3 (indicated as S3 in Figure 2). We did not include the probe that asked participants to order the events because of the complexities arising when comparing both the degree of change and the nature of this change. For questions involving multiple answers (such as the name of crash sites), we examined each answer separately. As illustrated in Figure 2, corrected responses were more common than either repeated or other responses, for both Surveys 2 and 3. Moreover, the proportions of corrected responses for event memories were greater on Surveys 2 and 3 than comparable figures for flashbulb memories. (All *t*-tests were significant at levels less than .005.) As for the uncorrected, inaccurate responses, on Survey 2, participants were as likely to say something else than repeat the errors they made on Survey 1. Survey 3 differed from Survey 2, in that for Survey 3, repeated responses were significantly greater than other responses, $t(390) = 3.74$, $p < .001$, $d = .38$. This latter finding suggests that, like flashbulb memories, a

stable event memory may be emerging somewhere between the first and third year. Importantly, though, the content of this stable memory differs from the content of flashbulb memories: Event memories are converging on an accurate rendering of the past, whereas flashbulb memories are converging on personally accepted and confidently held, even if inconsistent, renderings. This difference strengthens the interpretation we advanced when considering the predictor data: That the retention (and forgetting) of flashbulb and event memories over the long-term involves different processes.

Discussion

In this paper, we explored the long-term retention of flashbulb memories and the associated event memories. We investigated both the rate of forgetting and how different aspects of the memories might be forgotten at different rates, as well as the factors associated with the retention of both flashbulb memories and event memories.

Forgetting over the long-term

In the present study, the rate of forgetting of flashbulb memories in the first year was similar to that observed in other flashbulb memory studies, but, importantly, this rate of forgetting slowed substantially between the first and third year. This converges with other studies demonstrating a slowing rate of forgetting between the first and third years (Berntsen & Thomsen, 2005; Bohannon & Symons, 1992). The current study, however, has the advantage of using a test-retest paradigm. The present study brings into question the generality of the results of Schmolck et al (2002), who found an increased rate of forgetting between the first and third years using a similar test-retest paradigm. Horn (2001) has also questioned the generality of Schmolck et al., arguing that the steep forgetting Schmolck et al. observed between the fifteen-month and the thirty-second month tests may have been due to interference created by the announcement of the verdict of Simpson's civil trial in the sixteenth month. A similar confound probably does not arise for flashbulb memories of the terrorist attack of 9/11, the Challenger explosion, or the German invasion and withdrawal inasmuch as nothing so similar occurred during the retention periods of the respective studies for their respective samples.

When considered in conjunction with the other studies on long-term retention of flashbulb memories, then, the present study suggests that a slowing in the rate of forgetting after the first year is typical of flashbulb memories. In doing so, it supports and extends the findings of Talarico and Rubin (2003), that is, that the rate of forgetting follows a similar pattern as found for ordinary autobiographical memories. The rate of forgetting found in the present study was similar to that found in most diary studies: 20% or more the first year and between 5% to 10% thereafter.

Our results also suggest that different aspects of the flashbulb memories may be forgotten at different rates. Memory for emotions can be quite unreliable (Levine et al., 2006). Here we show that, despite the salience of the emotional reaction to flashbulb events such as 9/11, the memories of these emotional reactions tend to be forgotten more quickly than other aspects of the flashbulb memory, even over the long-term. The reason for this rapid forgetting needs to be further explored.

Our event memory results differ from those reported by Bohannon and Symons (1992) for the Challenger explosion, the only flashbulb memory study that examined event memory after three years. Bohannon and Symons found that forgetting continued to occur between the first and third year at approximately the same rate as it did in the first year, basing their conclusion on cross-sectional data, whereas we found a decline in the rate of forgetting when we examined longitudinal data, a finding consistent with work on memory for facts

(Baird, 1983, 1984; Baird, Baird, & Wittlinger, 1975). We explain this difference across studies below.

Do different factors influence flashbulb and event memories?

Although the pattern in the rate of forgetting was the same for flashbulb memories and event memories over a three-year period, subsequent analyses of our data suggested that different processes may underlie these similar patterns. In particular, we failed to find any relation between five predictors (residency, personal loss or inconvenience, emotionality, media attention, and ensuing conversation) and flashbulb memories, but found significant relations for four of the five (all but emotionality) predictors for event memories. Moreover, we found that inconsistent flashbulb memories reported on one survey were repeated on the next survey, whereas inaccurate event memories tended to be corrected on the next survey. These data were not at ceiling, limited in variability, nor specific to the coding scheme we employed.

Other studies of flashbulb memories of 9/11 using a test-retest procedure reported similar failures to find predictors for consistency (Curci & Luminet, 2006; Talarico & Rubin, 2003). It may be that various factors interact differently for different people. For example, some people may react emotionally and rehearse a flashbulb memory, whereas others may react emotionally and avoid rehearsing the memory. There could be enough variability in the population in the way various factors combine that comparing the effect of one of them on consistency across a population would be difficult. As a result, researchers might find low correlations for each factor in their flashbulb memory studies, even if in more controlled settings in which each factor is isolated they might find the predicted correlation. Luminet, Curci, and their colleagues have attempted to circumvent this problem by using structural equation modeling. Even this methodological advance has failed to produce uniform results (see Luminet, 2008, for a review).

Memory practices

Our finding that the levels of media attention and ensuing conversation are correlated with accuracy of event memories suggests these two variables should figure critically in any account of retention of event memories. Both of these activities increase the degree to which the event memory is rehearsed (Neisser, Winograd, Bergman, Schreiber, Palmer, & Weldon, 1996). In keeping with recent calls for a study of cultural products (Morling & Lamoreaux, 2008) and for viewing the mind as extended (Clark & Chalmers, 1998; Wilson, 2004), we also want to focus on the activity itself – the media attention or the ensuing conversation. Both media attention and ensuing conversation could be considered *memory practices of a community*, which refers to the way in which a community intentionally or unintentionally preserves its past (Bourdieu, 1977; for reviews, see Hirst & Manier, 2008; Hirst & Meksin, 2008; Zerubavel, 1997). For many social scientists, memorials and commemorations are the prototypical memory practices (Gillis, 1994). The practices of mass media in covering public, emotionally charged events and even the conversations people have about the event have also been treated as memory practices (Assmann & Czaplicka, 1995; Dayan & Katz, 1992; Hirst & Manier, 2008; Hirst & Echterhoff, 2008; Hoskins, 2007; Johnson, 2007). For ensuing conversation, people may be inclined to share memories of emotionally intense events, in part because they believe that doing so will help them deal with the emotion (Rimé, 2007a, 2007b). They may also discuss such events because of a social mandate. When meeting someone who has lost a loved one, some conversational acknowledgment of the death is mandatory. A similar social mandate may hold for public, emotionally charged events such as 9/11 (Mehl & Pennebaker, 2003). As we saw, the practice of ensuing conversation can vary across communities: The level of conversation in New York, for

instance, differed from the level of non-New Yorkers. Within a community, however, memory practices, such as conversations and media attention, appeared to be more uniform.

A social interactional approach suggests that a key to the difference between our results for event memory and those of Bohannon and Symons (1992) may lie in the memory practices surrounding 9/11 and the Challenger explosion. Although we cannot contrast the amount of conversation that followed the *Challenger* explosion with what followed the 9/11 attack, we can examine attention to the media, at least indirectly, by looking at media coverage (see Shapiro, 2006, for a similar analysis). Figure 3 plots the accuracy of event memory over a three-year period and a rough estimate of the amount of media coverage of the attack over the same time period.³ We estimated media coverage by using the *New York Times* as a reference text on Lexus-Nexus to determine the number of articles in the *Times* in which the phrase “September 11” or the conjunction of “Challenger” and “explosion” appeared. We then calculated the proportion of mentions over the number of days in the targeted period. To facilitate comparison, we recalibrated both the accuracy scores and the media proportions to z-scores calculated across the three time periods. As Figure 3 indicates, the pattern of forgetting Bohannon and Symons and we observed for event memory nicely mirrors the level of media attention. We found a similar pattern when we used *The Boston Globe* and *U.S. News and World Report* as reference texts in Lexus-Nexus.

Memory practices might also help account for the differences in changes in content we found for both flashbulb memories and event memories. Event memories tended to be corrected over time. Such correction is what one would expect if a community is constantly retelling the story of the attack. Unlike the retellings, by an individual, which may be subject to cumulative source monitoring failures (e.g., Johnson, 2006), a community retelling, especially in the media, tends to be fact-checked, and, *ceteris paribus*, is presumed to be “correct.” That is, the media can serve a social/cultural reality monitoring function (Johnson, 2007). The Michael Moore effect is a particularly vivid example of the ability of the media to correct inaccurate memories. The one exception to this general trend toward correction appears to be participants' memory for the names of the involved airlines. Memory for this fact continued to decline over the three-year period. This decline could also be explained by referring to community memory practices, if the names of airlines figured less critically in accounts of the 9/11 attack than, for instance, the number of planes. One could tell the story of 9/11 without mentioning the name of the airlines, but it would be much more difficult to avoid mentioning the locations of the attack or the number of planes involved. Films, such as *United 93*, came out after the last survey.

Whereas the memory practices of a community no doubt shaped the content of event memories of 9/11, there is little reason to expect that they should have a similar effect on flashbulb memories of 9/11. The memory practice of a community might lead its members to undertake similar mnemonic processing when it comes to memories for events, but, as we noted, members are more likely to be left to their own devices when it comes to flashbulb memories. Moreover, inasmuch as flashbulb memories are unique, a community as a whole rarely retells a single member's flashbulb memory across the community. Even when sharing of a flashbulb memory does occur in a small group, there may be no way to verify the accuracy of many details of a reported memory. Consequently, at least for objective features of flashbulb memories, the errors made at the end of the first year tended to persist into the third year. If a person falsely remembered that she was at work the first year, she tended to continue to remember (falsely) into the third year that she was at work. There may be nothing to lead her to suspect otherwise.

³Bohannon and Symons (1992) and we tested participants at slightly different delays. Hence, we discuss testing period rather specific testing delays.

In emphasizing the role of public memory practices to account for the shape of the forgetting of event memories, we have sought to understand what people remember and do not remember, not only in terms of individual internal cognitive processes, but also in terms of community activities. Both approaches are necessary to account for memory for public events. The different forgetting curves observed for event memories can be traced to the way the different events were covered by the media. Moreover, continuing media coverage can account for the corrections that took place over time with event memory. Clearly, an understanding of memories for public events such as the terrorist attack of 9/11 cannot be achieved by pointing only to internal cognitive processes or to social influences on memory, but to the interaction between the two.

Acknowledgments

We thank the many student coders without whom this research project would have been impossible, as well as Brett Sedgewick, who assisted with the supervision of this project.

Support from the James S. McDonnell Foundation and the National Institute of Mental Health, grant R01-MH0066972, is gratefully acknowledged.

References

- Assmann J, Czaplicka J. Collective memory and collective identity. *New German Critique* 1995;65:125–133.
- Baird, HP. The cognitive map of a city – 50 years of learning and memory. In: Bower, G., editor. *The psychology of learning and motivation: Advances in research and theory*. Vol. Vol 17. Academic Press; New York: 1983. p. 125-163.
- Baird HP. Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General* 1984;113:1–29. [PubMed: 6242406]
- Baird HP, Baird PO, Wittlinger RP. Fifty years of memory for names and faces: A cross-sectional study. *Journal of Experimental Psychology: General* 1975;104:54–75.
- Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Social and Personality Psychology* 1986;51:1173–1182.
- Baruch Y. Response rate in academic studies – A comparative analysis. *Human relations* 1999;52:421–439.
- Belli RF, Schuman H, Jackson B. Autobiographical misremembering: John Dean is not alone. *Applied Cognitive Psychology* 1997;11:187–209.
- Berntsen, D. Flashbulb memories and social identity. In: Luminet, O.; Curci, A.; Conway, M., editors. *Flashbulb memories: New issues and new perspectives*. Psychology Press; 2008. in press
- Berntsen D, Thomsen DK. Personal memories for remote historical events: Accuracy and clarity of flashbulb memories related to World War II. *Journal of Experimental Psychology: General* 2005;134:242–257. [PubMed: 15869348]
- Bohannon JN III. Flashbulb memories for the space shuttle disaster: A tale of two theories. *Cognition* 1988;29:179–196. [PubMed: 3168421]
- Bohannon, JN., III; Symons, VL. Flashbulb memories: Confidence, consistency, and quantity. In: Winograd, E.; Neisser, U., editors. *Affect and accuracy in recall*. Cambridge University Press; Cambridge: 1992. p. 65-90.
- Bourdieu, P. *Outline of a theory of practice*. Cambridge University Press; New York: 1977.
- Brown R, Kulik J. Flashbulb memories. *Cognition* 1977;5:73–79.
- Bruner, J. *Acts of Meaning*. Harvard University Press; Cambridge, MA: 1990.
- Budson AE, Simons JS, Sullivan AL, Beier JS, Solomon PR, Scinto LF, Daffner KR, Schacter DL. *Neuropsychology* 2004;18:315–327. [PubMed: 15099154]

- Budson AE, Simons JS, Waring JD, Sullivan AL, Hussoin T, Schacter DL. Memory for the September 11, 2001, terrorist attacks on year later in patients with Alzheimer's disease, patients with mild cognitive impairment, and healthy older adults. *Cortex* 2007;43:875–888. [PubMed: 17941346]
- Clark A, Chalmers D. The extended mind. *Analysis* 1998;58:10–23.
- Cohen. A power primer. *Psychological Bulletin* 1992;113:155–159. [PubMed: 19565683]
- Conway MA, Pleydell-Pearce CW. The construction of autobiographical memories in the self-memory system. *Psychological Review* 2000;107:261–288. [PubMed: 10789197]
- Curci A, Luminet O. Follow-up of a cross-national comparison of flashbulb and event memory for the September 11th attacks. *Memory* 2006;14:329–344. [PubMed: 16574589]
- Curci A, Luminet O, Finkenauer C, Gisle L. Flashbulb memories in social groups: A comparative test-retest study of the memory of French President Mitterand's death in a French and a Belgian group. *Memory* 2001;9:81–101. [PubMed: 11338939]
- Davidson PSR, Cook SP, Glisky EL. Flashbulb memories for September 11th can be preserved in older adults. *Aging, Neuropsychology, and Cognition* 2006;13:196–206.
- Dayan, D.; Katz, E. *Media events: The live broadcasting history*. Harvard University Press; Cambridge, MA: 1992.
- Ebbinghaus, H. *Memory; a contribution to experimental psychology*. Ruger, HA.; Bussenius, CE., translators. Dover; New York: 1913/1964. Original work published in 1913
- Echterhoff G, Hirst W. Thinking about memory for everyday and shocking events: Do people use ease-of-retrieval cues in memory judgments? *Memory & Cognition* 2006;34:763–775.
- Gillis, JR. *Commemorations: The politics of national identity*. Princeton University Press; Princeton, N.J.: 1994.
- Greenberg DL. President Bush's false 'flashbulb' memory of 9/11/01. *Applied Cognitive Psychology* 2004;18:363–370.
- Hirst W, Echterhoff G. Creating shared memories in conversation: Towards a psychology of collective memory. *Social Research* 2008;75:1071–1108.
- Hirst W, Manier D. Towards a psychology of collective memory. *Memory* 2008;16:183–200. [PubMed: 18324546]
- Hirst, W.; Meksin, R. A social-interactional approach to the retention of collective memories of flashbulb events. In: Luminet, O.; Curci, A.; Conway, M., editors. *Flashbulb memories: New issues and new perspectives*. Psychology Press; 2008. in press
- Horn DB. Confounding the effects of delay and interference on memory distortion: Commentary on Schmolck, Buffalo, and Squire. *Psychological Science* 2001;12:180. [PubMed: 11340930]
- Hoskins, A. *Media and memory*. Routledge; New York: 2007.
- Johnson MK. Memory and reality. *American Psychologist* 2006;61:760–771. [PubMed: 17115808]
- Johnson MK. Reality monitoring and the media. *Applied Cognitive Psychology* 2007;21:981–993.
- Kvavilashvili L, Mirani J, Schlagman S, Kornbrot DE. Comparing flashbulb memories of September 11 and the death of Princess Diana: Effect of time delays and nationality. *Applied Cognitive Psychology* 2003;17:1017–1031.
- Levine LJ, Prohaska V, Burgess SL, Rice JA, Laulhere RM. Remembering past emotions: the role of current appraisals. *Cognition and Emotions* 2001;15:292–417.
- Levine, L.J.; Safer, MA.; Lench, HC. Remembering and misremembering emotions. In: Sanna, L.J.; Chang, EC., editors. *Judgments of time: The interplay of thoughts, feelings, and behaviors*. Oxford University Press; NY: 2006. p. 271-290.
- Levine LJ, Whalen CK, Henker B, Jamner LD. Looking back on September 11, 2001: Appraised impact and memory for emotions in adolescents and adults. *Journal of Adolescent Research* 2005;20:497–523.
- Linton, M. Ways of searching and the contents of memory. In: Rubin, DC., editor. *Autobiographical memory*. Cambridge University Press; NY: 1986. p. 50-67.
- Luminet, O. Models for the formation of flashbulb memories. In: Luminet, O.; Curci, A.; Conway, M., editors. *Flashbulb memories: New issues and new perspectives*. Psychology Press; 2008. in press
- Luminet, O.; Curci, A.; Conway, M., editors. *Flashbulb memories: New issues and new perspectives*. Psychology Press; 2008. in press

- Luminet O, Curci A, Marsh E, Wessel I, Constantin T, Gencoz F. The cognitive, emotional, and social impacts of the September 11 attacks: Group differences in memory for the reception context and the determinants of flashbulb memory. *The Journal of General Psychology* 2004;17:171–181.
- McCloskey M, Wible CG, Cohen NJ. Is there a special flashbulb-memory mechanism? *Journal of Experimental Psychology: General* 1988;117:171–181.
- Mehl M, Pennebaker JW. The social dynamics of a cultural upheaval: Social interactions surrounding September 11, 2001. *Psychological Science* 2003;14:579–585. [PubMed: 14629689]
- Morling B, Lamoreaux M. Measuring culture outside the head: A meta-analysis of individual-collectivism in cultural products. *Personality and Social Psychology Review* 2008;12:199–221. [PubMed: 18544712]
- Neisser, U. Snapshots or benchmarks?. In: Neisser, U., editor. *Memory Observed: Remembering in Natural Contexts*. W. H. Freeman & Company; San Francisco: 1982. p. 43-48.
- Neisser, U.; Harsh, N. Phantom flashbulbs: False recollections of hearing the news about *Challenger*. In: Winograd, E.; Neisser, U., editors. *Affect and accuracy in recall*. Cambridge University Press; Cambridge: 1992. p. 9-31.
- Neisser U, Winograd E, Bergman ET, Schreiber CA, Palmer SE, Weldon MS. Remembering the earthquake: Direct experience vs. hearing the news. *Memory* 1996;4:337–357. [PubMed: 8817459]
- Olick J, Robbins J. Social memory studies: From “collective memory” to the historical sociology of mnemonic practices. *Annual Review of Sociology* 1998;24:105–140.
- Pezdek K. Event memory and autobiographical memory for events of September 11, 2001. *Applied Cognitive Psychology* 2003;17:1033–1045.
- Qin J, Mitchell KJ, Johnson MK, Krystal JH, Southwick SM, Rasmussen AM, Allen ES. Reactions to and memories for the September 11, 2001 terrorist attacks in adults with posttraumatic stress disorder. *Applied Cognitive Psychology* 2003;17:1081–1097.
- Rimé B. The social sharing of emotion as an interface between individual and collective processes in the construction of emotional climates. *Journal of Social Issues* 2007a;63:307–322.
- Rimé, B. Interpersonal emotion regulation. In: Gross, JJ., editor. *Handbook of emotion regulation*. Guilford Press; New York: 2007b. p. 446-485.
- Rubin, DC. Autobiographical memory tasks in cognitive research. In: Wenzel, A.; Rubin, DC., editors. *Cognitive methods and their application to clinical research*. American Psychological Association; Washington, D.C.: 2005. p. 219-241.
- Schmidt SR. Autobiographical memories of the September 11th attacks: Reconstructive errors and emotional impairment of memory. *Memory & Cognition* 2004;32:443–454.
- Scholck H, Buffalo EA, Squire LR. Memory distortions develop over time: Recollections of the O.J. Simpson trial verdict after 15 and 32 months. *Psychological Science* 2000;11:39–45. [PubMed: 11228841]
- Shapiro LR. Remembering September 11th: The role of retention interval and rehearsal on flashbulb and event memory. *Memory* 2006;14:129–147. [PubMed: 16484105]
- Sharot T, Martorella EA, Delgado MR, Phelps EA. How personal experience modulates the neural circuitry of memories of September 11. *Proceedings of the National Academy of Science* 2007;104:389–394.
- Smith MC, Bibi U, Sheard DE. Evidence for the differential impact of time and emotion on personal and event memories for September 11, 2001. *Applied Cognitive Psychology* 2003;17:1047–1055.
- Sobel, ME. Asymptotic confidence intervals for indirect effects in structural equation models. In: Leinhardt, S., editor. *Sociological methodology* 1982. Jossey-Bass; San Francisco: 1982. p. 290-312.
- Talarico JM, Rubin DC. Confidence, not consistency, characterizes flashbulb memories. *Psychological Science* 2003;14:455–461. [PubMed: 12930476]
- Talarico, JM.; Rubin, DC. Flashbulb memories result from ordinary memory processes and extraordinary event characteristics. In: Luminet, O.; Curci, A.; Conway, M., editors. *Flashbulb memories: New issues and new perspectives*. Psychology Press; 2008. in press
- Tekcan A, Berivan E, Gülgöz S, Er N. Autobiographical and event memory for 9/11: Changes across one year. *Applied Cognitive Psychology* 2003;17:1057–1066.

- Wagenaar WA. My memory: A study of autobiographical memory over six years. *Cognitive Psychology* 1986;18:225–252.
- Weaver CA. Do you need a “flash” to form a flashbulb memory? *Journal of Experimental Psychology: General* 1993;122:39–46.
- Wilson, RA. *Boundaries of the mind: The individual in the fragile sciences*. Cambridge University Press; New York: 2004.
- Winningham RG, Hyman IE, Dinnel DL. Flashbulb memories? The effects of when the initial memory report was obtained. *Memory* 2000;8:209–216. [PubMed: 10932791]
- Zerubavel, E. *Social mindscapes: An invitation to cognitive psychology*. Harvard University Press; Cambridge, MA: 1997.

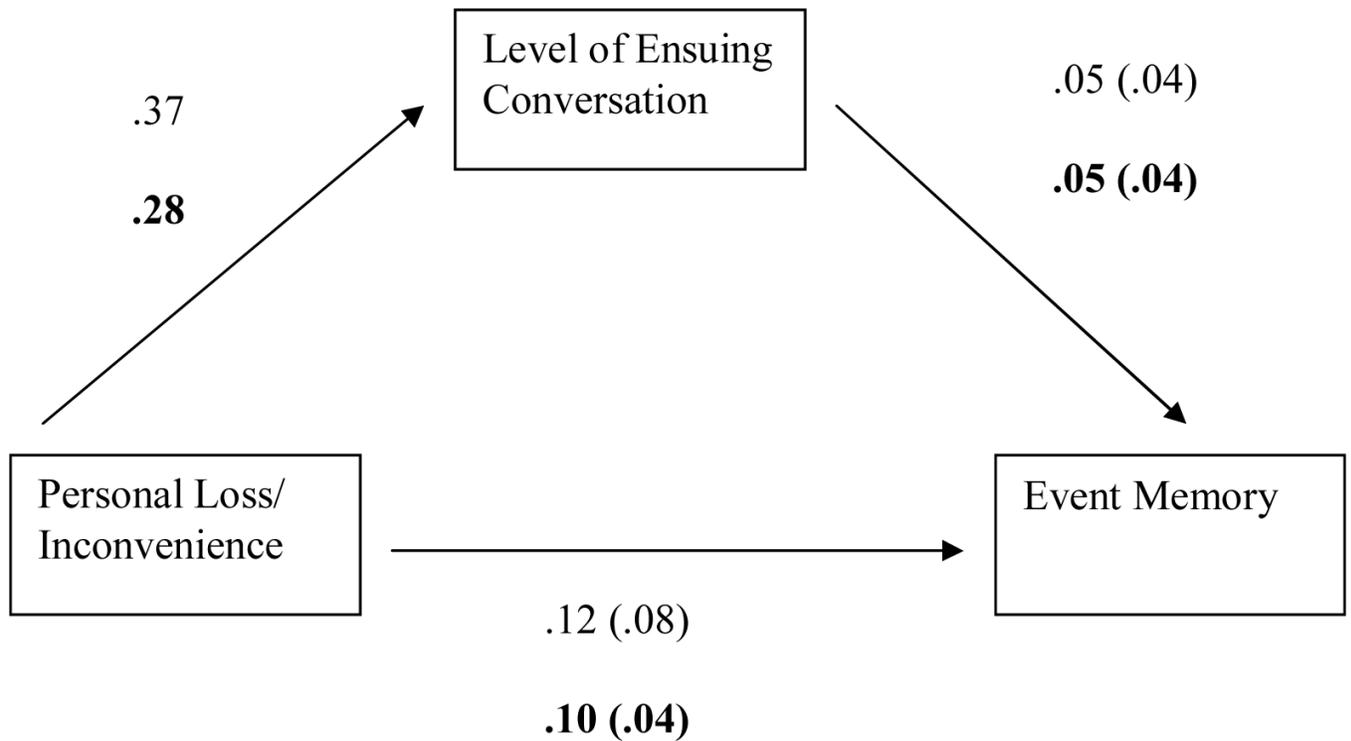


Figure 1. Mediation analysis following Baron & Kenny (1986)

A four step process with (1) regression coefficient from analysis with level of ensuing conversation as the dependent variable and personal loss as predictor, (2) regression coefficient from analysis with event memory as the dependent variable and level of ensuing conversation as the predictor, (3) regression coefficient from analysis with event memory as the dependent variable and personal loss as the predictor, and (4) regression coefficients from analysis with event memory as the dependent variable and both level of ensuing conversation and personal loss/inconvenience as predictors. These latter coefficients are in parentheses. Analysis for the 11-month period is in regular type. Analysis for 35-month period is in bold type. Partial mediation occurs if the regression coefficient for personal loss associated calculated in Step 4 is less than the related coefficient calculated in Step 3.

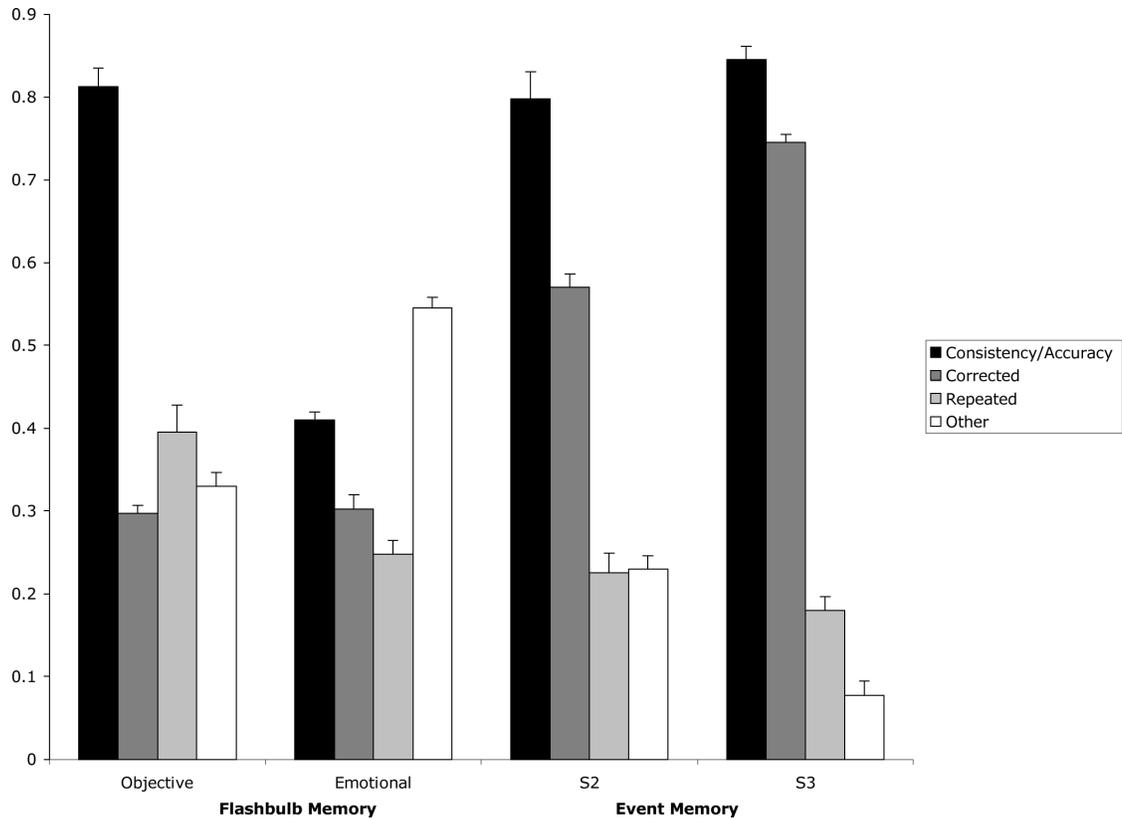


Figure 2.

For flashbulb memories, the proportion of consistent responses on Survey 2 that remained consistent on Survey 3 (as indicated by *consistency/accuracy*), as well as the proportion of inconsistent responses on Survey 2 *corrected*, *repeated*, or given another response on Survey 3. For event memory, the proportion of accurate responses on Survey 1 accurate on Survey 2 (S2: *consistency/accuracy*), the proportion of accurate responses on Survey 2 that were also accurate on Survey 3 (S3: *consistency/accuracy*), the proportion of inaccurate responses on Survey 1 *corrected*, *repeated* or given another response on Survey 2, and the proportion of inaccurate responses on Survey 2 *corrected*, *repeated* or given another response on Survey 3. Standard deviations are expressed in error bars.

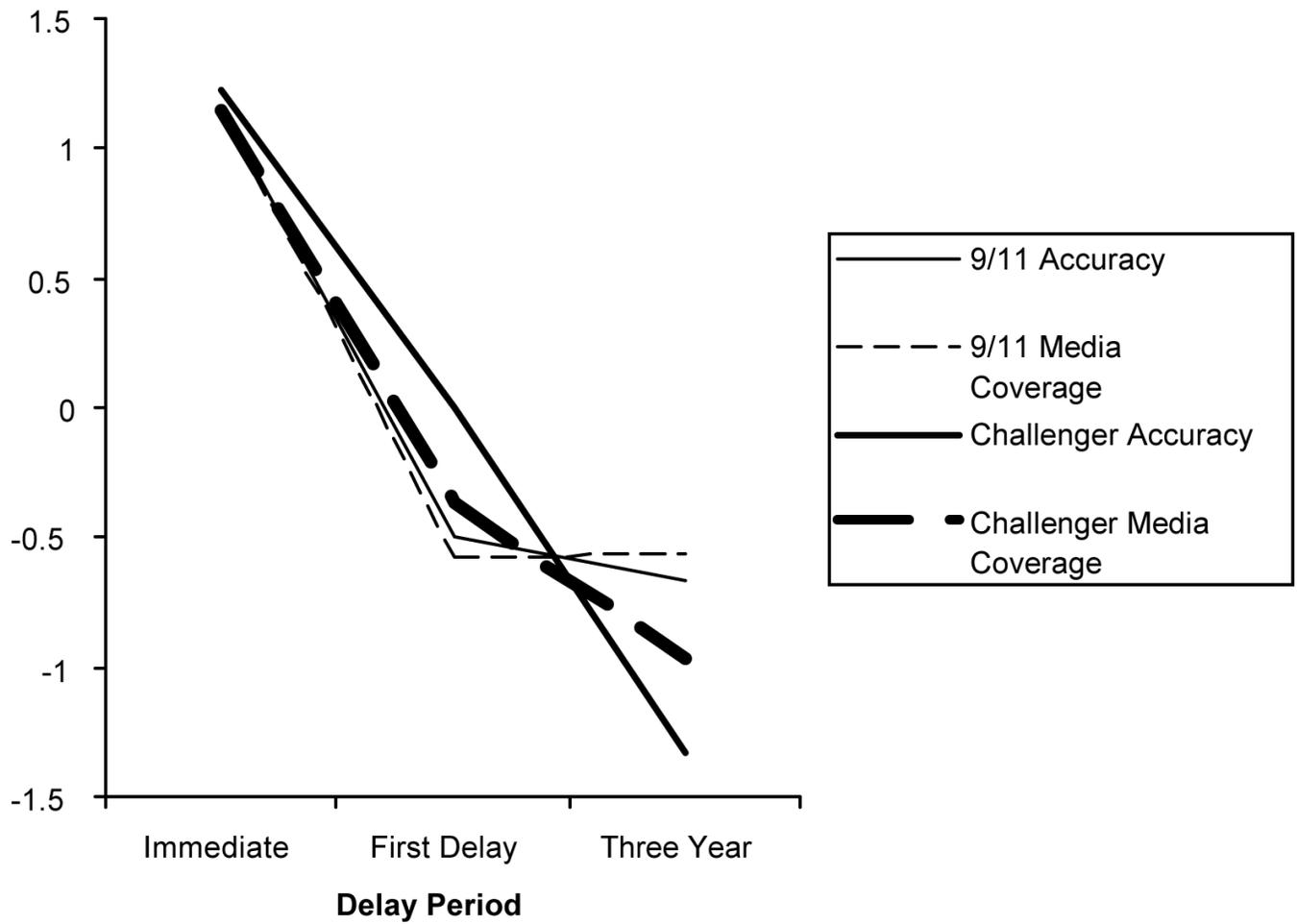


Figure 3. Relation between media coverage and memory accuracy for facts about the Challenger Explosion and the Terrorist Attack of September 11. Proportions of correctly remembered details were converted into z-scores, as were the frequencies with which an event was covered in the *New York Times*.

Table 1

Distribution of samples in which a participant responded on one or more than one survey

	Distribution Location							Total
	Bos	NH	NYC	DC	STL	Palo	SC	
Participants								
Multiple Responses								
S1, S2 & S3	-	28	168	55	96	24	20	391
S1 & S2	59	21	121	37	97	34	18	387
S1 & S3	-	1	40	12	10	5	4	72
S2 & S3	-	39	111	39	11	1	9	210
TOTAL	59	89	440	143	214	64	51	1060
Single Responses								
S1	15	152	572	212	117	128	71	1267
S2	-	74	158	104	33	32	64	465
S3	-	50	267	110	10	-	17	454
TOTAL	15	276	997	426	160	160	152	2186

Notes: Boston (Bos), New Haven (NH), New York City (NYC), Washington, DC (DC), St. Louis (STL), Palo Alto (Palo), Santa Cruz (SC).

Table 2

Relevant questions in Survey 1

FLASHBULB MEMORY

- 1) How did you first learn about it (what was the source of the information)?
- 2) Where were you?
- 3) What were you doing?
- 4) How did you feel when you first became aware of the attack?
- 5) Who was the first person with whom you communicated about the attack?
- 6) What were you doing immediately before you became aware of the attack?

EVENT MEMORY

- 7) How many airplanes were involved in the attack?
- 8) What airline or airlines had planes hijacked? How many from each airline?
- 9) In the vicinity of which cities did the airplanes end up?
- 10) Where was President Bush when the attack occurred?
- 11) Many people think that these are the most salient events that occurred in the attack:
 - a) The Pentagon was hit by a hijacked plane
 - b) A second World Trade Center Tower was hit by a hijacked plane
 - c) A World Trade Center Tower collapsed
 - d) A World Trade Center Tower was hit by a hijacked plane
 - e) A second World Trade Center Tower collapsed
 - f) A hijacked plane crashed outside of Pittsburgh

Please indicate the order in which you became aware of each event.

Please indicate the order in which the events actually occurred.

PREDICTORS

- 12) Did you suffer any personal losses in the attack? If so, please specify.
 - 13) Did the attack inconvenience your daily activities in some way? If so, please specify.
- For the following questions, we'd like you to tell us about your **CURRENT FEELINGS CONCERNING THE ATTACK**. Please indicate your response by marking the appropriate point on the scales provided. Note that you may indicate partial numbers (e.g. 3.5)
- 14) At this moment, how strongly or intensely do you feel **sad** about the attack?
 - 15) At this moment, how strongly or intensely do you feel **angry** about the attack?
 - 16) At this moment, how strongly or intensely do you feel **fear** about the attack?
 - 17) At this moment, how strongly or intensely do you feel **confusion** about the attack?
 - 18) At this moment, how strongly or intensely do you feel **frustration** about the attack?
 - 19) At this moment, how strongly or intensely do you feel **shock** about the attack?
 - 20) At this moment, what other emotions might you be experiencing? List and rate their intensity at this moment.
 - 21) How closely did you follow the media coverage? (Rate on a 1 – 5 scale.)
 - 22) How much have you talked about the attack since the announcement? (Rate on a 1 – 5 scale.)
 - 23) Since the attack please estimate the *percentage* of waking hours you have spent doing the following:
 - a. Watching television to get information about the attack _____
 - b. Talking to friends/relatives about the attack _____
 - c. Listening to radio coverage of the attack _____
 - d. Consulting the internet for news about the attack _____

- e. Reading published accounts of the attack _____
- f. Going about your daily activities _____
- g. Trying to help those who were hurt in some manner _____
- h. Walking around to see what you could see _____
- i. Other _____

DEMOGRAPHICS

- 24) Home of origin
 - 25) Permanent Residence
-

Table 3

Examples of coding schemes

*Coding for the probe "How did you first learn of the attack?" (Only code for **FIRST** response)*

(0) not stated, (1) TV, (2) Radio, (3) E-mail/IM, (4) Phone call (includes Phone messages), (5) Visual sighting, (6) Word-of-mouth, (7) Sounds/Screams/Sirens, (8) Other (Enter response in addendum)

Coding for the probe "Whom were you with when you learned of the attack?"

Inasmuch as more than one response was possible, coder had to indicate which of the following options were given on the survey.

Not stated, spouse/lover, child, sibling, other blood relative, parent-in-law, child-in-law, sibling-in-law, other in-law, close friend, acquaintance friend, colleague, roommate, teacher, student, classmate, neighbor, fellow commuter, stranger, government official (police, fire department, etc), medical personnel (doctor, nurse, EMT), alone, and other.

Table 4

Consistency and Confidence Ratings

	Survey 1 to 2		Survey 1 to 3	
	Mean	SD	Mean	SD
Overall Consistency	.63	.20	.57 ²	.23
Emotional Consistency	.42	.49	.37	.48
Correlation of Emotion z-scores	.48	.40	.42 ¹	.39
Overall Confidence ratings	4.41	.64	4.25	.93

¹ p < .05.

² p < .01.

Table 5

Facts accurately remembered: Means of the Accuracy Scores and Standard Deviations (in parentheses)

	Survey 1	Survey 2	Survey 3
Number of planes	.94	.86 ²	.81
Airline Names	.86 (.30)	.69 (.38) ²	.57 (.42) ²
Crash Sites	.93 (.19)	.92 (.20)	.88 (.25) ¹
Order of Events	.88 (.13)	.89(.11)	.86 (.14) ¹
Location of Bush	.87	.57 ²	.81 ²
Saw Moore's film	.87	.60 ²	.91 ²
Did not see film	.86	.54 ²	.71 ²
Overall	.88 (.14)	.77 (.21) ²	.78 (.23)

Notes. Data was either nominal or interval. Nominal data is reported as frequencies, interval data as proportions. The proportions are reported with standard deviations.

¹ p < .05.

² p < .01.

Table 6
 Mean consistency, confidence ratings, and accuracy as a function of residence and personal loss or inconvenience

	Residence		Personal loss/Inconvenience			
	Non-NYC	NYC	Full Sample		NYC	
	None	Present	None	Present	None	Present
Consistency						
S12	.65	.60 ¹	.64	.61	.64	.65
S13	.60	.56	.57	.56	.59	.57
Confidence						
S12	4.51	4.57	4.53	4.52	4.68	4.51
S13	4.21	4.32	4.23	4.27	4.18	4.11
Accuracy						
S1	.88	.89	.87	.90	.87	.89
S2	.74	.81 ²	.74	.81 ²	.70	.82 ²
S3	.75	.82 ¹	.77	.81 ¹	.72	.82 ¹

¹ p < .05.

² p < .01.

Table 7

Correlations between measures of emotion, media attention, ensuing conversation, and measures of consistency, confidence, and accuracy.

	Emotions			Media			Conversation		
	Average	Highest	S1	S2	S3	S1	S2	S3	
Consistency									
S12	-.10	.01	-.07	-.03	-.01	-.05	-.02	-.03	
S13	.02	.04	-.03	.03	.02	-.07	-.05	.01	
Confidence									
S12	.08	.10 ¹	.08	.13 ¹	.10 ¹	.18 ²	.18 ²	.04	
S13	.02	.08	.06	.04	.28 ²	.09	.05	.20 ²	
Accuracy									
S1	.05	.06	.19 ²	.04	.08	.16 ²	.02	.03	
S2	.03	.07	.21 ²	.11 ¹	.06	.19 ²	.18 ²	.02	
S3	-.01	.06	.14 ¹	.07	.12 ¹	.20 ²	.11 ¹	.12 ¹	

¹ p < .05.

² p < .01

Table 8

Average level of media attention and ensuing conversations for New Yorkers and non-New Yorkers with or without personal loss or inconvenience (PL/IN)

	Media			Conversations		
	S1	S12	S123	S1	S12	S123
New Yorkers						
Without PL/IN	4.36	4.01	3.78	4.53	4.12	3.89
With PL/IN	4.52	4.07	3.85	4.47	4.11	3.80
Overall	4.46	4.05	3.83	4.49	4.12	3.83
Non-New Yorkers						
Without PL/IN	4.27	3.82	3.65	4.25	3.74	3.27
With PL/IN	4.34	3.98	3.73	4.52 ²	4.11 ²	3.75 ²
Overall	4.29	3.87	3.67	4.33	3.85	3.56

¹ p < .05.

² p < .01