Flashbulb Memories for September 11th can be Preserved in Older Adults

Patrick S. R. Davidson, Shaun P. Cook, and Elizabeth L. Glisky
University of Arizona, Tucson, Arizona

Abstract

Flashbulb memories (FMs) are vivid, long-lasting memories for the source of surprising, arousing news. Laboratory studies have consistently found that older adults, especially those with below-average frontal lobe (FL) function, are impaired in source memory relative to young. We tested memory for the source of news concerning the September 11th terrorist attacks in older adults with high or low frontal lobe function and in young people. We found no age differences in source memory a year after the event and no differences related to FL function. Flashbulb memories may be different from usual cases of source memory because of their emotional content, personal importance, or social relevance.

Flashbulb memories (FMs) are vivid recollections of the source of news about an emotionally arousing, shocking event (Brown & Kulik, 1977) (for reviews, see Conway, 1995; Pezdek, 2003). Flashbulb memories are generally more durable than memories for everyday experiences (e.g., Christianson, 1989; Davidson & Glisky, 2002a), and there has been considerable disagreement about whether FMs depend on the same cognitive and neural mechanisms as less emotional memories. If they do, then one might expect that FMs, like other source memories, might be particularly susceptible to aging and declining frontal lobe (FL) function.

On the face of it, studies of FM in the real world appear to be very similar to research on source memory (broadly defined, see Johnson et al., 1993) in the laboratory—in both cases participants must recollect when, where, or from whom they learned information. In the laboratory, source memory declines reliably with aging (for reviews, see Spencer & Raz, 1995; Prull et al., 2000). Furthermore, source memory problems occur primarily in older adults with below-average FL function, as indicated by neuropsychological testing (e.g., Craik et al., 1990; Henkel, Johnson, & De Leonardis, 1998; Schacter et al., 1991). For example, research in our lab has shown that older people with low FL function are poorer than young people at remembering the voice that said a word, the list in which a word was presented, or the room in which a chair appeared, whereas older adults with high FL function perform like young (Davidson & Glisky, 2002b; Glisky et al., 1995; Glisky et al., 2001).

In FM situations, however, findings in older adults have not always followed this pattern, although results are inconsistent. First, although aging is usually accompanied by a significant decline in source memory in the laboratory, this has not always been the case for FM. In an initial study, Davidson and Glisky (2002a) asked young and older people to report how they learned of the deaths of Diana, Princess of Wales, and Mother Teresa of Calcutta. The older people were just as likely as the young to remember where, when, and from whom they learned this information (see also Gerdy et al., 2004; Otani et al., 2004; Wright et al., 1998). In an
earlier study, however, Cohen, et al. (1994) reported an age-related decline in memory for hearing about the resignation of Prime Minister Thatcher in England (see also Tekcan & Peynircioglu, 2002). Second, although in laboratory studies FL function has usually been related to source memory in older adults (e.g., Craik et al., 1990; Davidson & Glisky, 2002b; Glisky et al., 1995; 2001; Henkel et al., 1998; Schacter et al., 1991), there was no such relation in our real world study of FM (Davidson & Glisky, 2002a): Surprisingly, older people with below average FL function retained slightly more about how they learned of the deaths of Princess Diana and Mother Teresa than those with above-average FL function. We speculated that the emotional arousal associated with FM may have enhanced the FL processes important for the encoding of source information, leading to better source memory than normally expected in the older adults.

The goal of the present study was to replicate our previous report (Davidson & Glisky, 2002a). We used a double assessment procedure to measure participants' retention of information over time: Soon after the September 11th terrorist attacks on New York City and Washington D.C., and then approximately 1 year later, participants were asked how they heard the news. We tested young people and older adults who were divided on the basis of FL function using a method similar to our laboratory studies of source memory (Glisky et al., 1995; 2001). People were also asked to rate several factors related to their encoding and rehearsal of the event, which have been related to FM in previous studies (see Conway, 1995). We hypothesized that if FM is just another instance of source memory, then it should be impaired in older adults, especially in those with low FL function, relative to the young. If, however, FM is special in some way, then these groups may not differ. We tested a larger sample than in our previous study, in order to increase statistical power and the likelihood of uncovering significant group differences. Also, although it could be argued that the events featured in our previous study (the deaths of Princess Diana and Mother Teresa) were not sufficiently emotional to qualify as true FM's, that objection seems not to apply here.

**Method**

**Participants**

Forty-six young adults (mean age 22 years, mean education 14 years) and 43 older adults (mean age 74 years, mean education 15 years) were each tested on two occasions, approximately 1 year apart. The young adults were recruited from undergraduate psychology classes at the University of Arizona, and either volunteered or were awarded course credit for completing the initial test, and were offered $6 for completing the retest. The older adults were recruited from our laboratory's pool of community-dwelling older people, and were free of depression and dementia. They had previously completed neuropsychological screening, which was used to create a composite score thought to reflect FL function for each individual. The composite was uncovered in an earlier factor analysis of a normative group of 100 older adults, after removing variance in performance associated with age (see also Glisky et al., 1995; 2001). It consisted of: the number of categories achieved on a modified Wisconsin Card Sorting Test (Hart et al., 1988), number of words produced in a phonemic fluency test (Spreen & Benton, 1977), mental arithmetic from the Wechsler Adult Intelligence Scale—Revised (WAIS-R) (Wechsler, 1981), and backward digit span and mental control from the Wechsler Memory Scale—III (WMS-III) (1997). These measures were used to create an average $z$ score describing FL function. Older adults were categorized as high or low in FL function, depending on whether their $z$ scores were above or below the normative group mean.

Table 1 shows the demographic and neuropsychological characteristics of the older adults. Nineteen scored above the mean on the FL factor, and 24 were below the mean on it. Separate independent sample $t$ tests showed that, although the two groups were similar in age ($t < 1$), the high FL group had higher levels of education, $t (41) = 2.49, p = .02$, slightly higher Mini
Mental State Exam scores (MMSE) (Folstein, Folstein, & McHugh, 1975), \( t(41) = 2.40, p = .02 \), and higher vocabulary scores, \( t(41) = 3.44, p = .001 \), than the low FL group.

### Materials

The questionnaire was similar to those used in previous studies (e.g., Cohen et al., 1994; Davidson & Glisky, 2002a; Neisser & Harsch, 1992; Pillemer, 1984). Questions concerned place (i.e., *Where were you?*), the participants' interrupted activity, the participants' ensuing activity, other people's presence, time of day, and the source of the news (i.e., *How did you hear the news, i.e., from what source?*). Self ratings (from 1 = very low to 5 = very high) of encoding (surprise, emotional intensity, the event's importance to the participant and to the world), and rehearsal (frequency of thoughts about, media exposure to, and discussion of the event) were also collected.

We also asked questions about a comparison autobiographical memory (the most interesting event from the few days previous to September 11, 2001), which were analogous to those above, with the exception that this section did not include a specific question about source (i.e., it did not ask *How did you hear about this interesting event?*).

### Procedure

Shortly after September 11, 2001 (i.e., between 3 and 21 days), participants completed the first questionnaire. Older adults were interviewed by telephone, were read the questions one at a time, and dictated their answers to us. Responses were recorded as close to verbatim as possible. Young adults were tested either individually or in groups of approximately 20; in both cases they wrote down their responses on a printed questionnaire. This difference in the procedure was necessary in order to collect initial reports as soon as possible after the event. All respondents, however, were encouraged to provide as detailed answers as possible. All participants gave informed consent to take part in the project; their rights were protected, and our institution’s human subjects research guidelines were followed.

The surprise retest was completed approximately 1 year (i.e., between 11 and 13 months) after the initial session, by telephone. The second questionnaire was identical to the first, with the exception of the comparison event section. For the comparison event, participants were first asked whether they could remember the most interesting event they had experienced in the few days previous to the terrorist attacks. If they could not, or recalled a different event from the one on the initial test, they were provided with a minimal cue (e.g., “a party” or “a movie”) to help increase their chance of recalling the previously reported event.

### Scoring

On the initial test, participants were assigned a score of 0 if they did not provide an answer to a question, or a 1 if they did; therefore the maximum possible score for the FM event was 6; for the comparison memory it was 5. Because of the different number of questions for each event, scores were converted to proportions.

On the delayed test, the consistency of each participant’s responses over time was measured in a way similar to previous studies (e.g., Cohen et al., 1994; Davidson & Glisky, 2002a; Neisser & Harsch, 1992). Two naïve judges rated how similar answers on the initial test were to those on the delayed test, using a scale of 0, 1, or 2 for each question. Judges assigned a 0 if an answer was not given on the delayed test or if it was incongruent with the initial test, 1 if the answer on the delayed test was somewhat consistent with or less specific than the initial test (e.g., “in my wife’s car, turning off Elm Street” at the initial test, compared to “in a car” at retest), or 2 if it was very similar between tests. The raters agreed on 82% of the questions (in cases of disagreement the median rating was used). If a participant did not answer a question on the...
initial test, then that item was omitted from the congruence scoring. For each event, a summed congruence score was derived, based on a maximum of six questions for the flashbulb event, and five questions for the comparison event, and converted to a proportion, to reflect how much information was retained over the 1-year delay. Ratings of the encoding and rehearsal variables were selected from the initial test.

Results

Retention of FM

Overall proportions of FM recalled on the initial test and retrieved after the delay are shown in Table 2a. All three groups scored at the ceiling in initial memory (Ms = 1.00 for each group). A year later, all three groups retained approximately 75–80% of the source information initially reported, and all individuals were able to remember at least some of the source details. A between groups ANOVA (young, high FL, low FL) showed no reliable difference among groups (F < 1.00). Also, when we computed the correlation between FL function and flashbulb memory retention in the older adults, it was not reliable (r = -.04) (Wright, 2003). We also looked at retention of answers to the individual questions, which ranged from 71% (for what people did afterwards) to 86% (for who they were with), and are shown in Table 2b. Separate 3 (group) X 6 (question) mixed ANOVAs revealed no reliable differences among groups, and no interactions between group and question (p > .05). Finally, we examined whether FM was correlated with a composite measure of medial temporal lobe function, based on standard tests of item memory,1 but there was no significant relation (r = .20).

Encoding and Rehearsal Variables

The three groups did differ in their ratings of two of the FM encoding variables and one measure of rehearsal, shown in Table 3. Separate one-way ANOVAs and subsequent Bonferroni t-tests indicated that, first, there was a difference in emotional intensity, F (2, 85) = 3.75, MSE = 1.13, p = .03: The low FL group (M = 4.63) reported significantly greater emotional intensity than the young group (M = 3.96; p = .04). Second, there was a similar pattern in personal importance, F (2, 85) = 5.19, MSE = .85, p = .007, with the low FL group (M = 4.62) indicating a greater level than the young (M = 3.89; p = .007). Third, the amount of media coverage differed among groups, F (2, 85) = 3.29, MSE = .92, p = .04, again with the low FL group (M = 4.58) experiencing more than the young (M = 3.98; p = .04). Importantly, however, no significant differences were found between the ratings of the high and low FL groups.

Retention of the Comparison Event

Unlike memory for the FM event, memory for the comparison event was quite poor: many participants could retain nothing at all about it after the 1-year delay, despite each person being prompted with a specific cue. Only 42% of the older adults were able to retain any information at all about this event, compared to 65% of the young adults, a difference which was statistically significant, χ² [1, N = 86] = 4.46, p = .04. There was little difference between the FL factor groups in the likelihood of remembering the comparison event.

Discussion

Memory for how one learned about the September 11th terrorist attacks was examined in young people and two groups of older adults. People retained more than 75% of this information over 1 year, a rate comparable to other FM studies using this procedure (e.g., Cohen et al., 1994).

---

1The composite measure of medial temporal lobe function (see Glisky et al., 1995, 2001) was based on Logical Memory I, Faces I, and Verbal Paired Associates I from the WMS-III, Visual Paired Associates II from the Wechsler Memory Scale—Revised (WMS-R) (Wechsler, 1987), and Long Delay Cued Recall on the CVLT (Delis et al., 1987).
Older adults performed just as well as young people, and the low FL older adults were just as likely to retain FM details as the high FL group. These findings are consistent with some previous reports (Davidson & Glisky, 2002a; Gerdy et al., 2004; Otani et al., 2004; Wright et al., 1998), but inconsistent with others (Cohen et al., 1994; Tekcan & Peynircioglu, 2002). In particular, both of the latter studies reported reduced FM in older people, and Cohen et al. speculated that the age deficit might be attributable to reduced frontal function in older adults. Our findings suggest that this is not the case, and that older adults with low FL function are just as likely as others to have FMs.

What might account for the discrepant findings across studies of FM and aging? One possibility concerns statistical power. In the present study, power was relatively low, and so it is possible that aging and/or low FL function could have a small, but reliable, effect on FM, which we simply failed to detect here. However, in the present study the groups performed so similarly to one another (e.g., FM retention Ms = .78 and .77 in the young and low FL aged groups, respectively) that any reliable differences uncovered with larger groups would likely have been very small, and of little practical importance. Also, even with a relatively large number of young and aged participants (n = 65 in each group), Gerdy et al. (2004) still reported age equivalence in FM retention for September 11th.

Another possibility is that differences in encoding and rehearsal influence whether age effects are seen in FM. For example, in Tekcan and Peynircioglu's study (2002), the young adults reported greater amounts of rehearsal than the old people. Rehearsal is generally associated with greater retention of FM information (for a review see Conway, 1995; but see Neisser & Harsch, 1992), so this may account for the observed age differences in FM in that report. In the present study, the FM event was rated as more emotionally arousing, more personally important, and more often seen in media coverage by the low FL older adults than the young people. These factors may have obscured a possible difference between these groups in FM. Yet the low FL and high FL groups of older people did not differ on any of the encoding or rehearsal variables, and also showed equivalent levels of FM retention. As well, in a previous study Davidson and Glisky (2002a) found that even when young and older adults made similar ratings of encoding and rehearsal variables, the groups were also equivalent on FM.

The lack of influence of aging and reduced FL function on flashbulb memory appears to stand in contrast to memory for less emotional, everyday events. For example, in our previous study, we found that memory for another personal experience (e.g., attending a party, or seeing a film) that occurred around the same time as the flashbulb event was correlated with our neuropsychological measures of FL and MTL function in older adults. Similarly, laboratory studies have usually shown that older adults are impaired in source memory (see Spencer & Raz, 1995; Prull et al., 2000), especially when they have low FL function (Davidson & Glisky, 2002b; Glisky et al., 1995, 2001). In the present study, few older adults retained much about a personal experience from around the same time as September 11th, making it difficult to examine in depth.

Overall, people in our study retained less information about the comparison event than the FM event (as in Christianson (1989) and Davidson & Glisky (2002a)). Although this contrast was not the main focus of our study, it is different from a recent finding by Talarico and Rubin (2003), who actually reported better retention of an everyday event than a FM event. It is unclear what might account for the difference between our studies and theirs, but there were several methodological differences that could provide an explanation. For example, retention in the Talarico and Rubin study was intentional, and people were instructed to generate their own descriptive cues to retrieve the event in the future. Overall their levels of retention across a 7-month interval appear lower (approximately 65% of original details retained) than found in many previous FM studies, including ours (75–80% retention).
Why might FM be unaffected by aging and reduced FL function, when other kinds of source memory are significantly impaired? We speculated previously that the emotional arousal associated with FM situations may boost or enhance in some way the FL processes involved in the encoding of source information (Davidson & Glisky, 2002a). In a similar vein, Rahhal et al. (2002) reported no age differences in source memory when people were instructed to focus on emotional or socially relevant information about the sources (e.g., whether one of two speakers was truthful). Elevated levels of emotional arousal, social relevance, or personal importance associated with FMs may increase attentional processes such that contextual and source information are more likely to be encoded and integrated with the target event (e.g., Kahneman, 1973; Revelle & Loftus, 1992). Further study is thus required to assess the differences between flashbulb memory in the real world and source memory in the laboratory. To this end, we are currently investigating source memory for emotional events in the laboratory in young and older adults. Other studies are planned to explore how source memory is affected by such variables as personal importance and rehearsal, which are likely associated more with everyday than laboratory events.

The present report, along with Davidson and Glisky (2002a), has potentially identified boundary conditions on the influences of aging and FL function on source memory in older adults, and is consistent with other recent suggestions that age-related decline in episodic memory can be attenuated when people are presented with emotional materials (Carstensen & Turk-Charles, 1994), or are instructed to focus on emotional aspects of stimuli (Rahhal et al., 2002). Generalizations based on the present data, however, should be made cautiously because of the drawbacks associated with studies of memory for real-world events, such as FMs. Such studies are based on retrospective report, with little experimental control over variables (e.g., individual differences in reaction to the news) that may influence performance. As well, the context in which each person learns of a flashbulb event is unique to the individual and unknown to us, requiring measures of consistency of participants' responses over time rather than accuracy. For these reasons, among others, it seems critical to continue studies of emotion, source memory, and aging in a more controlled laboratory setting.

Acknowledgements

This research was supported by fellowships from the Natural Sciences and Engineering Research Council of Canada and the Canadian Institutes of Health Research to P. D., and a National Institute on Aging grant (AG 14792) to E. G. A preliminary version of this report received the 2003 Completed Doctoral Research Award from the Division of Adult Development and Aging of the American Psychological Association. We thank Claire Conlon, Maria Marquine, Michael Robinson, and Andrea Soulé for assistance. We are grateful to Lee Ryan and Malcolm Binns for discussion of the statistical analyses.

References


Spreen, O.; Benton, AL. Neurosensory Center Comprehensive Examination for Aphasia (NCCEA). Revised. Victoria, British Columbia, Canada: University of Victoria Neuropsychology Laboratory; 1977.


Wright DB. Making friends with your data: Improving how statistics are conducted and reported. British Journal of Educational Psychology 2003;73:123–136. [PubMed: 12639281]

Table 1
Characteristics of Older Adults Participating in Immediate and 1-Year Delayed Test as a Function of Group

<table>
<thead>
<tr>
<th></th>
<th>High FL (n = 19)</th>
<th>Low FL (n = 24)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Age (years)</td>
<td>74.58</td>
<td>4.54</td>
</tr>
<tr>
<td>Education (years)</td>
<td>16.42</td>
<td>3.10</td>
</tr>
<tr>
<td>MMSE (/30)</td>
<td>29.32</td>
<td>1.06</td>
</tr>
<tr>
<td>Vocabulary(^a)</td>
<td>71.63</td>
<td>5.59</td>
</tr>
<tr>
<td>FL function score(^b)</td>
<td>.54</td>
<td>.32</td>
</tr>
</tbody>
</table>

\(^a\) From the Wechsler Abbreviated Scale of Intelligence (1999).

\(^b\) See text for explanation of score.
<table>
<thead>
<tr>
<th>Test</th>
<th>Young M</th>
<th>SEM</th>
<th>High FL M</th>
<th>SEM</th>
<th>Low FL M</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immediate</td>
<td>1.00</td>
<td>.00</td>
<td>1.00</td>
<td>.00</td>
<td>1.00</td>
<td>.00</td>
</tr>
<tr>
<td>1-year delay</td>
<td>.78</td>
<td>.03</td>
<td>.75</td>
<td>.04</td>
<td>.77</td>
<td>.03</td>
</tr>
<tr>
<td>Test</td>
<td>Young M</td>
<td>SEM</td>
<td>Young M</td>
<td>SEM</td>
<td>High FL M</td>
<td>SEM</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------</td>
<td>-----</td>
<td>-----------</td>
<td>-----</td>
<td>-----------</td>
<td>-----</td>
</tr>
<tr>
<td>Place</td>
<td>.83</td>
<td>.04</td>
<td>.86</td>
<td>.05</td>
<td>.82</td>
<td>.05</td>
</tr>
<tr>
<td>Interrupted activity</td>
<td>.77</td>
<td>.05</td>
<td>.72</td>
<td>.05</td>
<td>.71</td>
<td>.07</td>
</tr>
<tr>
<td>Ensuing activity</td>
<td>.72</td>
<td>.05</td>
<td>.74</td>
<td>.08</td>
<td>.67</td>
<td>.08</td>
</tr>
<tr>
<td>Other people</td>
<td>.83</td>
<td>.05</td>
<td>.89</td>
<td>.08</td>
<td>.88</td>
<td>.08</td>
</tr>
<tr>
<td>Time of day</td>
<td>.86</td>
<td>.04</td>
<td>.70</td>
<td>.09</td>
<td>.71</td>
<td>.09</td>
</tr>
<tr>
<td>Source of news</td>
<td>.69</td>
<td>.06</td>
<td>.64</td>
<td>.10</td>
<td>.84</td>
<td>.08</td>
</tr>
</tbody>
</table>

Table 2B
Proportion of Each Element of FM Retained as a Function of Group
<table>
<thead>
<tr>
<th></th>
<th>Young</th>
<th></th>
<th></th>
<th>High FL</th>
<th></th>
<th></th>
<th>Low FL</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SEM</td>
<td>M</td>
<td>SEM</td>
<td>M</td>
<td>SEM</td>
<td>M</td>
<td>SEM</td>
</tr>
<tr>
<td>Surprise</td>
<td>4.70</td>
<td>.12</td>
<td>4.95</td>
<td>.05</td>
<td>4.88</td>
<td>.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional intensity</td>
<td>3.96</td>
<td>.17</td>
<td>4.50</td>
<td>.26</td>
<td>4.63</td>
<td>.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal importance</td>
<td>3.89</td>
<td>.14</td>
<td>4.00</td>
<td>.25</td>
<td>4.62</td>
<td>.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>World importance</td>
<td>4.96</td>
<td>.03</td>
<td>4.84</td>
<td>.09</td>
<td>4.88</td>
<td>.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thought</td>
<td>3.93</td>
<td>.12</td>
<td>4.16</td>
<td>.18</td>
<td>4.08</td>
<td>.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Media exposure</td>
<td>3.98</td>
<td>.16</td>
<td>4.05</td>
<td>.21</td>
<td>4.58</td>
<td>.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discussion</td>
<td>3.29</td>
<td>.13</td>
<td>3.42</td>
<td>.25</td>
<td>3.33</td>
<td>.23</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>