American Psychological Association
Psychology and Aging

ISSN: 0882-7974
2020, Vol. 35, No. 4, 459–472
http://dx.doi.org/10.1037/pag0000467

Age-Related Differences in Flashbulb Memories: A Meta-Analysis

Sarah J. Kopp, Laura E. Sockol, and Kristi S. Multhaup
Davidson College

Recent meta-analyses reveal age-related declines in short-term memory (STM), working memory, associative memory, prospective memory, face memory, recognition, and recall. The present meta-analyses extend this work beyond predominantly laboratory-based tasks to a naturalistic phenomenon. *Flashbulb memories* are vivid autobiographical recollections for the circumstances in which one learns of a distinct event that may be surprising, emotional, or personally important (the *reception event*). The existing literature on aging and flashbulb memories includes inconsistent findings. The present meta-analyses included 16 studies ($N = 1898$) that examined flashbulb memory in nonclinical samples of younger adults (below age 40 years) and older adults (above age 60 years). Findings, after exclusion of an outlier, suggest a small-to-moderate age-related impairment in flashbulb memory scores ($k = 14$, Hedges' $g = -0.30$, 95% CI [-0.45, -0.15], $p < .001$) that was not moderated by study characteristics. After exclusion of an outlier, older adults' flashbulb memories were also significantly less consistent across time than younger adults' ($k = 7$, Hedges' $g = -0.29$, 95% CI [-0.47, -0.11], $p = .002$). Secondary analyses investigated age-related differences in the presence and consistency of canonical categories of flashbulb memories and encoding and rehearsal variables associated with flashbulb memory formation and retention. Age-related differences were found only for consistency of memory for ongoing activity at the time of the reception event, favoring younger adults ($k = 3$, Hedges' $g = -0.40$, 95% CI [-0.65, -0.15], $p = .002$). Overall, these findings are consistent with age-related impairment in flashbulb memory formation and retention.

Keywords: aging, flashbulb memory, meta-analysis, autobiographical memory, activity memory

Supplemental materials: http://dx.doi.org/10.1037/pag0000467.supp

Meta-analyses on predominantly laboratory-based memory tasks reveal age-related differences favoring younger adults. An early meta-analysis (Verhaeghen, Marcoen, & Goossens, 1993) found that older adults score lower than younger adults on tests of STM (cf. Bopp & Verhaeghen, 2005), working memory, prose recall, paired-associate recall, and list recall. Similar age-related declines are reported on tests of associative memory (Old & Naveh-Benjamin, 2008; see also Spencer & Raz, 1995, regarding context memory), focal and nonfocal prospective memory (Kliegel, Jäger, & Phillips, 2008; Uttl, 2011), and many measures of face memory (Martschuk & Sporer, 2018). Additional meta-analyses show the same pattern for recognition (Fraundorf, Hourihan, Peters, & Benjamin, 2019) and recall (Rhodes, Greene, & Naveh-Benjamin, 2019). In these nine meta-analyses, the only reported age-related difference favoring older adults was memory of older adults' faces, and it was a marginal effect (Martschuk & Sporer, 2018).

In their discussion of age-related memory impairment on multiple measures, however, Verhaeghen et al. (1993) make a critical point: Finding that older adults perform less well than younger adults on carefully constructed laboratory tasks does not necessarily translate to older adults having deficits in everyday memory performance that impair functioning (see also West, 1986). Moreover, prior work has suggested that different cognitive processes underlie older adults' performance on typical laboratory- and simulated everyday-memory tasks (Kempe, Kalicinski, & Memmert, 2015). In addition, the link between neuropsychological test performance, which includes memory performance, and everyday memory performance is inconsistent (McAlister & Schmitter-Edgecombe, 2016; Ossher, Flegal, & Lustig, 2013). Furthermore, the age-related memory impairments typically observed on laboratory-based tasks can be eliminated when naturalistic materials are used. For example, in a study of recall for price information, although younger adults demonstrated greater recall when associations between prices and items were arbitrary, there was no difference in the performance of older and younger adults when the task was implemented using realistic prices for grocery items

This article was published Online First April 20, 2020.
Sarah J. Kopp, Laura E. Sockol, and Kristi S. Multhaup, Department of Psychology, Davidson College.

Sarah J. Kopp is now at the University of Cincinnati College of Medicine.

This research was supported by a Davidson Research Initiative (Duke Endowment) grant to Sarah J. Kopp. Portions of an initial meta-analysis were presented on November 14–17, 2019, at the 2019 Annual Meeting of the Psychonomic Society in Montreal, Quebec, Canada. The data we used for analysis can be accessed at https://osf.io/cjgwx/. We thank Serena Hu for her coding assistance and Maurya M. Boyd for helpful discussions. Manuscripts were scored according to a coding manual (available at https://osf.io/34jvz/).

Correspondence concerning this article should be addressed to Kristi S. Multhaup, Department of Psychology, Davidson College, Box 7000, Davidson, NC 28035-7000. E-mail: krmulthaup@davidson.edu

(Castel, 2005). The present research extends the prior meta-analyses on aging and laboratory-based memory performance to a naturalistic memory phenomenon, namely, the presence of vivid memories for learning important news such as where one was when they learned that JFK had been shot and how one learned about the 9/11 attacks.

Flashbulb Memories

Flashbulb memories are exceptionally vivid, but not necessarily accurate recollections of the circumstances in which one learns of a significant event (e.g., Brown & Kulik, 1977; Cohen, Conway, & Maylor, 1994; Hirst & Phelps, 2016). They are a specific subset of autobiographical memories that encode contextual information about the *reception event*, that is, when an individual first receives news of a major event, rather than factual details of the event itself (Cohen et al., 1994). Although most of this literature examines flashbulb memories for public events, such as the terrorist attacks on September 11, 2001 (e.g., Hirst et al., 2015), flashbulb memories can also form from personal events, such as learning of the death of a loved one (Demiray & Freund, 2015). In short, these are the kinds of memories that people share in everyday conversations. To lay the foundation for the dependent measures included in the present meta-analysis of aging and this type of naturalistic memory, we summarize key points from the flashbulb memory literature.

When sharing flashbulb memories, people often report seemingly insignificant details. For example, 14 years after President Kennedy's assassination, participants reported minutiae such as the weather of the day or the fact that they dropped a carton of Viceroy cigarettes upon learning the news (Brown & Kulik, 1977). Such detailed reports reminded Brown and Kulik (1977) of the flash before a photograph is taken; thus, they coined the term *flashbulb memories*. In a point discussed less often in subsequent research, the authors also emphasized that flashbulb memories are unlike photographs in that they are incomplete records of a moment in time, whereas photographs completely capture the details present in a moment (p. 75).

Although some details of a moment may be lost, Brown and Kulik (1977) use the term *canonical categories* to denote common aspects of the reception event reported by individuals during free recall of learning important news. Their original canonical categories included *informant* (source of news), *place* (location of news reception), *ongoing activity* (action that was interrupted by news), *aftermath* (activities that followed news reception), *own affect* (emotional state of self), and *others' affect* (emotional state of others). A review of the literature from 1977 to 2010 (Kızılöz & Tekcan, 2013) found that most flashbulb memory studies assessed the canonical categories of source, place, and ongoing activity. In addition, most studies also included assessments of memory for the *time* of the reception event and *others present* when the individual learned of the event. The present meta-analyses examine age-related differences in the presence and consistency of memory for five canonical categories: source (we use this term which includes TV, e.g., rather than *informant* which suggests a person), location (synonymous with *place*), ongoing activity, time, and others present.

Research has explored the event characteristics and behaviors associated with flashbulb memory formation and retention (e.g., Brown & Kulik, 1977; Conway et al., 1994; Er, 2003; Finkenauer et al., 1998). *Encoding variables,* such as the personal importance of the event or the intensity of emotion, are characteristics of the reception event which are associated with flashbulb memory formation (Davidson & Glisky, 2002). Although various models of flashbulb memory formation emphasize different encoding variables (e.g., Luminet, 2018), four variables emerge as common predictors of flashbulb memories: (1) the individual's level of investment in the event, (2) the importance of the event, (3) the degree to which the event produced surprise, and (4) the emotional intensity of the individual's response during the reception event. In addition, *rehearsal variables*, such as the degree to which individuals think about (covert rehearsal) or discuss (overt rehearsal) the event, are believed to be crucial in maintaining flashbulb memories (Davidson & Glisky, 2002). Thus, the present meta-analyses also examine age-related differences in these five encoding and rehearsal variables.

Aging and Flashbulb Memories

The current literature assessing age-related differences in flashbulb memories includes inconsistent findings. Some studies report that flashbulb memories are impervious to age-related effects (Conway, Skitka, Hemmerich, & Kershaw, 2009; Davidson, Cook, & Glisky, 2006; Kvavilashvili, Mirani, Schlagman, Erskine, & Kornbrot, 2010; Wolters & Goudsmit, 2005; Wright, Gaskell, & O'Muircheartaigh, 1998). Others, however, find evidence of an age-related impairment in flashbulb memories (e.g., Cohen et al., 1994; Kensinger, Krendl, & Corkin, 2006; Tekcan & Peynircioğlu, 2002; Yarmey & Bull, 1978). We report the first meta-analyses of studies investigating age-related differences in the everyday memory phenomenon of flashbulb memories. The present research examines age-related differences in (1) the presence of flashbulb memories, (2) memory for canonical categories of reception events, (3) consistency of flashbulb memory scores across time, (4) consistency of canonical category scores across time, and (5) encoding and rehearsal variables associated with flashbulb memory formation and retention.

Method

Search Procedure and Study Selection

A flowchart depicting the search process and review of studies for eligibility is shown in Figure 1. Studies were identified through a search of the PsycINFO database through January 2019 using the following key terms: *flashbulb memory* AND *age*, *flashbulb memory* AND *older adults*, *flashbulb memory* AND *elderly*, *emotional memory* AND *aging*, *vivid memory* (with participant age 65+ as a qualifier), and *flashbulb memory* (with participant age 65+ as a qualifier). Additional studies were identified through review of references of eligible manuscripts, contacting authors of reviewed studies to identify unpublished and in-press work, and review of abstracts from the 2018 annual meeting of the Psychonomic Society. After removal of duplicates, this process identified 189 manuscripts for review. Full-text manuscripts were reviewed by the primary investigator to determine eligibility. To be eligible for inclusion in the meta-analyses, studies were required to meet the following criteria:

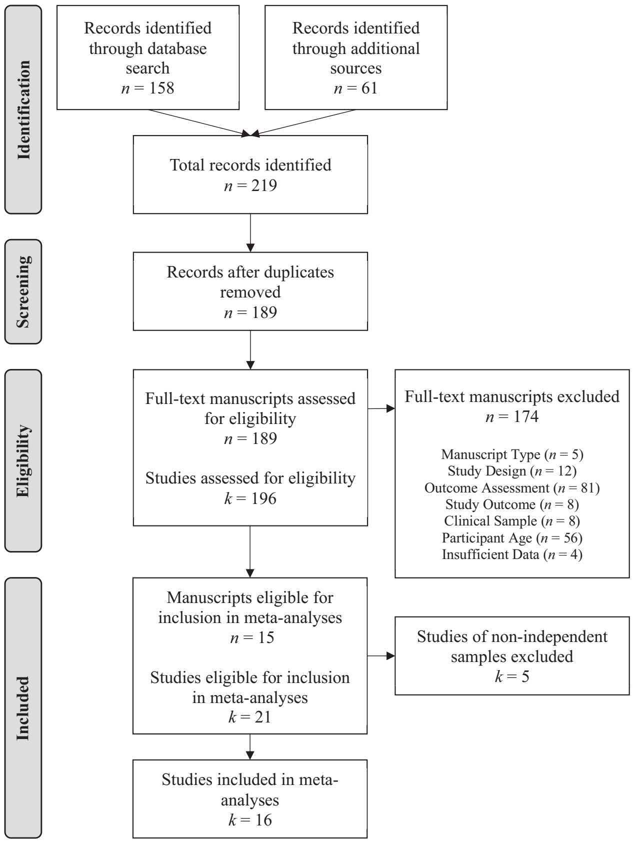

Figure 1. Flowchart illustrating identification of included studies. Results of 21 eligible studies were reported in 15 different manuscripts. Results of one study were reported in two manuscripts; effect sizes for this study were calculated using the manuscript reporting the most comprehensive data. Three manuscripts reported results from multiple studies. When a manuscript reported results from more than one study using the same sample, in order to meet the assumption of independence of effect sizes, results from the study with the smallest effect size were included in the meta-analyses.

Manuscript type. The manuscript presents an empirical study of human subjects.

Study design. The study design is cross-sectional or cross-sequential. Case studies were excluded, as these studies report insufficient data for the calculation of effect size(s).

Outcome assessment. The study reports at least one objective measure of flashbulb memory. Studies that assessed age-related differences on nonflashbulb memory outcomes, such as memory for emotional pictures, were excluded.

Study outcome. The study assessed the difference in at least one measure of flashbulb memory between older and younger adults.

Nonclinical sample. Participants were recruited from a nonclinical population (i.e., no reported difficulties in memory or cognitive functioning). Studies of clinical populations, such as older adults with cognitive impairments, were excluded.

Participant age. Average age of participants in the older adult group was over 60 years, and average age of participants in the younger adult group was below 40 years.

Available data. The article reports sufficient data to calculate effect size(s) representing the difference in flashbulb memories (score and/or consistency measure) between an older and younger group of participants. When necessary, authors were contacted to request additional data necessary for the calculation of effect size(s). In one case, the primary investigator and a second coder used the tool WebPlotDigitizer (Version 4.2; Rohatgi, 2019) to extract data for the calculation of effect size from a published figure.

Review of full-text manuscripts identified 15 manuscripts reporting results from 21 studies that met full inclusion criteria. Results of one study were reported in two manuscripts (Kvavilashvili et al., 2010; Kvavilashvili, Mirani, Schlagman, Wellsted, & Kornbrot, 2009); effect sizes for this study were calculated using the manuscript reporting the most comprehensive data for each analysis. One manuscript (Kvavilashvili et al., 2009) included three independent samples so all samples were retained. Two manuscripts reported results from multiple studies (Davidson & Glisky, 2002, reported two studies; Tekcan et al., in press, reported five studies). For these latter two manuscripts, in order to meet the assumption of independence of effect sizes, results from the study with the smallest effect size were included in the meta-analyses. Thus, results from 15 manuscripts, representing 16 independent studies, were included in the meta-analyses.

Coding of Included Studies

Manuscripts were scored according to a coding manual (available at https://osf.io/34jvz/). All eligible studies were independently coded by the primary investigator and a second coder; interrater reliability was high (94.9% agreement). Disagreement was resolved by consensus, with final determination made by the third author as necessary.

Effect sizes. Effect sizes were calculated to represent the difference in flashbulb memory scores between older and younger adult participants using the standardized mean difference (Hedges' *g*), which is less susceptible to upward bias than Cohen's *d* (Hedges, 1981; Lipsey & Wilson, 2001). Effect sizes were coded so that positive effect sizes indicate superior memory (participants remembered more characteristics of the reception event or had greater consistency in memories for the reception event between assessments) for older adults compared with younger adults. Similarly, for analyses of encoding and rehearsal variables, positive effect sizes indicate higher scores for older adults compared with younger adults (e.g., greater investment, more overt rehearsal).

In our primary analyses, we evaluated two types of effect sizes for flashbulb memories. The first represents the difference in overall flashbulb memory scores between older and younger adults at the initial assessment. Most studies ($n = 13$, 87%) reported this score as a composite of multiple items assessing the presence or level of detail of elements of the flashbulb memory event. In two studies, this score was derived from a single item: Yarmey and Bull (1978) assessed flashbulb memory with a single item (coded yes or no) indicating whether participants could remember what they were doing at the time of the reception event, and Greene, Loftus, Grady, and Levine (2018) assessed the clarity and vividness of participants' memory for the reception event with a nine-point Likert scale.

The second effect size represents the difference in the consistency of older and younger adults' memories between the initial assessment and a follow-up assessment. This was typically scored following a procedure developed by Neisser and Harsch (1992): Participants respond to a series of questions regarding elements of the reception event (e.g., location, others present) after the event and at a subsequent assessment point; responses at the two assessment points are then compared and receive full scores for identical responses, partial scores for similar or slightly distorted responses, and no score for discrepant or absent responses. The sum of participants' scores on the elements of the memory for the reception event represents the overall consistency of participants' flashbulb memories.

In addition to calculating effect sizes representing (1) the age-related difference in overall flashbulb memory scores and (2) their consistency over time, we also calculated effect sizes representing (3) age-related differences in five canonical categories of flashbulb memories (source, location, ongoing activity, others present, and time of the reception event) and (4) their consistency over time. Finally, we calculated effect sizes representing (5) age-related differences in encoding and rehearsal variables associated with flashbulb memory formation and retention: investment, importance, surprise, emotionality, and overt rehearsal.

In order to meet the assumption of independence of effect sizes, a single effect size for each outcome was calculated for each independent study included in the meta-analyses. If a study reported more than one measure for a given outcome, the outcome measure most closely aligned with the construct was used to calculate the effect size. For example, if a study reported both general interest in politics and knowledge of government, effect sizes for investment were calculated using scores on general interest in politics. Similarly, if a study reported both personal importance of the event and national importance of the event, scores on personal importance of the event were used to calculate the effect size. When studies reported more than one rehearsal-related outcome, the variable most closely related to speaking about the reception event was used to calculate the effect size for overt rehearsal.

Study characteristics. In addition to extracting data for the calculation of effect size(s), coders scored all studies on characteristics of the study and demographic characteristics of the sample.

Study characteristics included the country in which the study was conducted, study design (cross-sectional vs. cross-sequential), the flashbulb memory event, and timing of assessments (delay interval and retention interval; defined below). Study design was coded as cross-sectional if participants were assessed at a single point in time and cross-sequential if participants were assessed at multiple time points. *Delay interval* was coded as the number of days between the flashbulb memory event and the initial assessment. In cross-sequential studies, *retention interval* was coded as the number of days between the initial and follow-up assessment. In one study (Kvavilashvili et al., 2010), half of the participants completed an interim assessment two weeks after the initial assessment, with all participants completing a final assessment approximately two years after the flashbulb event. Effect sizes for this study were calculated using information from the full sample at the initial and final assessments.

Demographic characteristics of the samples included age (range, mean, and standard deviation for older and younger adults), gender (proportion of female participants in older and younger adult groups), and race/ethnicity (proportion of racial/ethnic minority participants in older and younger adult groups).

Methodological quality. Coders scored methodological quality on an 8-point scale, which included components of study design quality and comprehensiveness of reporting of sample characteristics, with higher scores representing more rigorous methodological quality. Studies received up to one point each for four elements of study design quality: naïve raters, measuring consistency of flashbulb memories, screening for cognitive impairment, and delay interval. Studies with a delay interval of less than one week received one point, studies with a delay interval between one week and one month received 0.5 points, and studies with a delay interval greater than one month received zero points. Studies also received up to one point each for reporting four demographic characteristics of included samples: gender, age, education level, and race/ethnicity. For reporting age, studies received one point for reporting the mean, standard deviation, and range for age for both older and younger adult groups, 0.67 points for reporting two elements (e.g., mean and standard deviation but not range) for both groups, and 0.33 points for reporting one element (e.g., range only) for both groups.

Data Analysis

Separate meta-analyses were conducted for all outcomes assessed in three or more independent studies. All meta-analyses were conducted using random effects models (Borenstein, Hedges, Higgins, & Rothstein, 2009) using Comprehensive Meta-Analysis Version 3.3.070 (Borenstein, Hedges, Higgins, & Rothstein, 2017).

Following recommendations to assess outliers in meta-analysis using a combination of quantitative approaches with visual tools (Aguinis, Gottfredson, & Joo, 2013; Huffcutt & Arthur, 1995), outliers were identified by assessment of the sample-adjusted meta-analytic deviancy statistic (SAMD) followed by visual assessment of scree plots. To maximize the variance available for moderator analyses, only studies with absolute SAMD values greater than 2.24 which were also discrepant with the visual distribution of the rank-ordered scree plot of SAMD values were excluded from subsequent analyses (Aguinis et al., 2013).

Heterogeneity of effect sizes was estimated using Cochran's Q statistic and I^2 values. Cochran's Q tests for the presence of heterogeneity; a significant value indicates that the variability in the effect sizes between studies is greater than can be accounted for by within-study variability (Cochran, 1954). The I^2 statistic estimates the proportion of the observed variance in effect sizes that cannot be attributed to chance; values of 25%, 50%, and 75% indicate low, moderate, and high degrees of heterogeneity, respectively (Higgins, Thompson, Deeks, & Altman, 2003).

Publication bias was assessed through visual examination of funnel plots, fail-safe N values, and Duval and Tweedie's (2000a, 2000b) trim-and-fill procedure. First, effect sizes for each study were plotted against the study's standard error, and these funnel plots were evaluated for asymmetry. An asymmetric distribution suggests the presence of publication bias (Sterne et al., 2011). Rosenthal's (1979) fail-safe N estimates the number of missing studies necessary to result in a nonsignificant overall effect size; we used the recommended rule for a fail-safe N exceeding the threshold of $5k + 10$, where k is the number of included studies, to indicate an unlikely number of unpublished studies. Duval and Tweedie's (2000a, 2000b) trim-and-fill procedure was used to evaluate potential publication bias and, when missing studies were identified, to provide an unbiased estimate of the true effect size.

When the Q and I^2 values indicated significant heterogeneity among effect sizes, we conducted exploratory moderator analyses to determine if characteristics of studies were systematically associated with effect sizes for the primary outcomes (flashbulb memory scores and consistency of flashbulb memory scores). One categorical characteristic (study design) and five continuous characteristics (methodological quality, delay interval, retention interval, and mean age of older and younger adults) were assessed as potential moderators. We conducted categorical moderator analyses for subgroups including three or more independent studies using mixed-effects analysis of variance (ANOVA). Continuous moderators were assessed using meta-regression; separate meta-regressions were conducted for each moderator.

Results

Study Characteristics

Sixteen studies, representing 1,898 total participants, were eligible for inclusion in the meta-analyses. Characteristics of included studies are reported in Table 1. Studies were primarily conducted in English-speaking countries ($k = 11$, 69%), including the United States ($k = 6$), United Kingdom ($k = 4$), Canada ($k = 1$), and Ireland ($k = 1$). One study (Yarmey & Bull, 1978) included participants from both the United States and Canada. The remaining studies were conducted in Turkey ($k = 2$), Germany ($k = 1$), Japan ($k = 1$), and the Netherlands ($k = 1$). Half of the included studies used a cross-sequential design ($k = 8$, 50%), seven used cross-sectional designs (44%), and the methodological design of one study (Tekcan et al., in press) was unknown. Studies assessed a wide range of flashbulb events; the most common event was the September 11th terrorist attack on the United States ($k = 5$, 31%).

Characteristics of samples and assessments for the included studies are reported in Table 2. The mean age for older adults ranged from 66.88 to 74.58 years, with an average of 71.15 years ($SD = 2.54$). The mean age for younger adults ranged from 19.18 to 37.72 years, with an average of 26.37 years ($SD = 5.98$). Delay interval ranged from 5 to 4,745 days; the mean delay interval was 1081.50 days ($SD = 1706.40$, $Mdn = 26$). Among cross-sequential studies, the retention interval ranged from 61 to 717 days ($M = 299.75$, $SD = 202.28$, $Mdn = 283$).

Characteristics of the methodological quality of the included studies are presented in Table 3. Methodological quality scores ranged from 0.00 to 5.67, with a mean of 3.37 ($SD = 1.58$). The quality of the design of included studies was generally low, although in fairness, the detailed reporting of participant demographics has only recently been encouraged and that comprised half of the total score. Only three studies (19%) reported that coders were naïve to the study hypothesis and/or age group of participants. Despite the literature suggesting flashbulb memories are subject to distortion over time, only half of the included studies ($k = 8$, 50%) measured consistency of flashbulb memories. Less than half of the included studies ($k = 7$, 44%) reported screening for cognitive impairment. Three studies (19%) conducted the initial assessment within one week of the event, with an additional five studies (31%) conducting the initial assessment between 1 week and 1 month following the event. Studies provided limited information regarding demographic characteristics of the included samples. Although most studies provided information regarding participant gender (63%), age (94%), and level of education (56%), only six studies (38%) provided information regarding all these characteristics, and no studies reported information regarding participant race/ethnicity.

Age-Related Differences in Flashbulb Memories

Fifteen studies assessed age-related differences in flashbulb memories. Six studies found that flashbulb memory scores for older adults were significantly lower than for younger adults, and

Table 1
Characteristics of Included Studies

Study	Country	Study design	Event
Bohn and Berntsen (2007)	Germany	CS	Fall of Berlin Wall
Cohen, Conway, and Maylor (1994)	United Kingdom	CQ	Resignation of Margaret Thatcher
Davidson, Cook, and Glisky (2006)	United States	CQ	September 11, 2001 terrorist attacks
Davidson and Glisky (2002) Study 2	United States	CQ	Death of Mother Theresa
Denver, Lane, and Cherry (2010)	United States	CS	September 11, 2001 terrorist attacks
Gerdy, Multhaup, and Ivey (2007)	United States	CQ	September 11, 2001 terrorist attacks
Greene, Loftus, Grady, and Levine (2018)	Ireland	CQ	May 2018 abortion referendum
Kensinger, Krendl, and Corkin (2006)	United States	CQ	Explosion of Columbia Shuttle
Kvavilashvili, Mirani, Schlagman, Wellsted, and Kornbrot (2009), Study 1	United Kingdom	CS	Death of Princess Diana
Kvavilashvili et al. (2009) Study 2	United Kingdom	CS	Death of Princess Diana
Kvavilashvili et al. (2009) Study 3	United Kingdom	CQ	September 11, 2001 terrorist attacks
Otani et al. (2005)	Japan	CQ	Nuclear accident in Ibaraki
Tekcan et al. (in press), Study 1	Turkey	NR	Challenger shuttle explosion
Tekcan and Peynircioğlu (2002)	Turkey	CS	Death of President Ozal
Wolters and Goudsmit (2005)	Netherlands	CS	September 11, 2001 terrorist attacks
Yarmey and Bull (1978)	United States and Canada	CS	Assassination of John Fitzgerald Kennedy

Note. CS = cross-sectional; CQ = cross-sequential; NR = not reported.

Table 2
Characteristics of Samples and Assessments for Included Studies

Study	*N*	Older adults		Younger adults		Delay interval (days)	Retention interval (days)
		n	*M* age	*n*	*M* age		
Bohn and Berntsen (2007)	133	94	69.38	39	34.41	4745	NA
Cohen et al. (1994)	120	60	71.60	60	22.40	12	342
Davidson et al. (2006)	65	19	74.58	46	22.00	12	366
Davidson and Glisky (2002), Study 2	49	29	73.97	20	19.87	21	183
Denver et al. (2010)	58	19	73.70	39	23.30	NR	NA
Gerdy et al. (2007)	151	76	73.95	75	19.18	5	140
Greene et al. (2018)	115	17	66.88	98	31.43	31	61
Kensinger et al. (2006)	110	62	72.60	48	24.80	7	224
Kvavilashvili et al. (2009), Study 1	183	94	67.30	89	37.72	2190	NA
Kvavilashvili et al. (2009), Study 2	85	38	67.61	47	34.21	2555	NA
Kvavilashvili et al. (2009), Study 3	164	80	71.43	84	32.05	6	717
Otani et al. (2005)	63	9	69.56	54	24.09	21	365
Tekcan et al. (in press), Study 1	172	78	72.67	94	24.99	NR	NR
Tekcan and Peynircioğlu (2002)	106	47	71.53	59	20.29	1095	NA
Wolters and Goudsmit (2005)	54	20	70.50	34	24.80	61	NA
Yarmey and Bull (1978)	270	44	NR	226	NR	4380	NA
Overall *M*	118.63	49.13	71.15	69.50	26.37	1081.50	299.75
Overall *SD*	59.59	29.08	2.54	47.65	5.98	1706.40	202.28

Note. NA = not assessed; NR = not reported.

one study found the opposite pattern, with Hedges' *g* ranging from −0.73 to 0.66 (see Table 4; for a forest plot presenting effect sizes for individual studies, see the online supplemental materials). The overall effect size indicated that older adults' flashbulb memory scores were significantly lower than younger adults', with an effect size in the small range (Hedges' $g = -0.25$, 95% CI [−0.42, −0.07], $p = .005$). The SAMD values for four studies were greater than 2.24. Visual examination of the scree plot indicated that values for three studies were consistent with the overall distribution (Kensinger et al., 2006; Tekcan et al., in press; Yarmey & Bull, 1978); these studies were retained in subsequent analyses. The value for one study (Greene et al., 2018) was discrepant with the overall distribution. After removal of this outlier, the overall effect size remained significant (Hedges' $g = -0.30$, 95% CI [−0.45, −0.15], $p < .001$). The Q statistic indicated significant heterogeneity among effect sizes, and the I^2 value of 49 suggested a medium degree of heterogeneity.

There was no evidence of publication bias in the included studies. The funnel plot appeared symmetrical, and the trim-and-fill procedure suggested no missing studies. The fail-safe N value of 99 exceeded the tolerance value (80) so it is unlikely that the reported age-related difference is due to large numbers of studies

Table 3
Methodological Quality of Included Studies

Study	Method. quality	Study design quality				Comprehensiveness of reporting			
		Naïve raters	Assess consist.	Cog. screen	Delay interval	Gender	Age	Education	Race/ethnicity
Bohn and Berntsen (2007)	3.00	NR	0.00	0.00	0.00	1.00	1.00	1.00	0.00
Cohen et al. (1994)	4.17	NR	1.00	1.00	0.50	1.00	0.67	0.00	0.00
Davidson et al. (2006)	4.83	1.00	1.00	1.00	0.50	0.00	0.33	1.00	0.00
Davidson and Glisky (2002), Study 2	4.83	1.00	1.00	1.00	0.50	0.00	0.33	1.00	0.00
Denver et al. (2010)	0.67	NR	0.00	0.00	NR	0.00	0.67	0.00	0.00
Gerdy et al. (2007)	4.00	NR	1.00	0.00	1.00	1.00	1.00	0.00	0.00
Greene et al. (2018)	4.50	NR	1.00	0.00	0.50	1.00	1.00	1.00	0.00
Kensinger et al. (2006)	4.67	1.00	1.00	1.00	1.00	0.00	0.67	0.00	0.00
Kvavilashvili et al. (2009), Study 1	3.67	NR	0.00	1.00	0.00	1.00	0.67	1.00	0.00
Kvavilashvili et al. (2009), Study 2	3.67	NR	0.00	1.00	0.00	1.00	0.67	1.00	0.00
Kvavilashvili et al. (2009), Study 3	5.67	NR	1.00	1.00	1.00	1.00	0.67	1.00	0.00
Otani et al. (2005)	3.50	NR	1.00	0.00	0.50	1.00	1.00	0.00	0.00
Tekcan et al. (in press), Study 1	2.00	NR	NR	NR	NR	0.00	1.00	1.00	0.00
Tekcan and Peynircioğlu (2002)	3.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	0.00
Wolters and Goudsmit (2005)	1.67	NR	0.00	0.00	0.00	1.00	0.67	0.00	0.00
Yarmey and Bull (1978)	0.00	NR	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note. Method. quality = methodological quality; Assess consist. = measure consistency of flashbulb memory from Time 1 to Time 2; Cog. screen = screen participants for cognitive impairment; NR = not reported.

Table 4
Random Weighted Effect Sizes From Meta-Analysis of Flashbulb Memory Scores

Study	Hedges' *g*	95% CI Lower bound	95% CI Upper bound	*p*	SAMD
Bohn and Berntsen (2007)	−0.14	−0.54	0.25	.485	0.59
Cohen et al. (1994)	−0.01	−0.51	0.50	.982	1.34
Davidson et al. (2006)	0.00	−0.53	0.53	>.999	1.01
Davidson and Glisky (2002), Study 2	−0.59*	−1.16	−0.02	.044	−1.24
Denver et al. (2010)	−0.13	−0.67	0.41	.632	0.44
Gerdy et al. (2007)	−0.50**	−0.84	−0.15	.005	−1.48
Greene et al. (2018)	0.66*	0.15	1.18	.012	4.94[a]
Kensinger et al. (2006)	−0.73***	−1.11	−0.34	<.001	−2.60
Kvavilashvili et al. (2009), Study 1	−0.50**	−0.80	−0.21	.001	−1.79
Kvavilashvili et al. (2009), Study 2	0.02	−0.40	0.45	.914	1.28
Kvavilashvili et al. (2009), Study 3	−0.18	−0.48	0.13	.254	0.45
Tekcan et al. (in press), Study 1	0.11	−0.19	0.40	.492	2.39
Tekcan and Peynircioğlu (2002)	−0.49*	−0.88	−0.10	.013	−1.30
Wolters and Goudsmit (2005)	−0.45	−1.00	0.10	.108	−0.79
Yarmey and Bull (1978)	−0.55**	−0.87	−0.22	.001	−2.44

	k	Hedges' *g*	95% CI Lower bound	95% CI Upper bound	*p*	*Q* (*df*)	I^2
Total (all studies)	15	−0.25**	−0.42	−0.07	.005	38.47 (14)***	64
Total (outlier excluded)	14	−0.30***	−0.45	−0.15	<.001	25.61 (13)*	49

Note. SAMD = sample-adjusted meta-analytic deviancy statistic.
[a] Outlier excluded.
* $p < .05$. ** $p < .01$. *** $p < .001$.

that found no age-related difference existing in unpublished files rather than the literature.

As both the Q statistic and I^2 value suggested significant heterogeneity among effect sizes, we assessed six characteristics of studies as potential moderators of effect size: study design, methodological quality, delay interval, retention interval, the mean age of older adults, and the mean age of younger adults. We conducted a mixed-effects ANOVA to evaluate whether the difference between older and younger adults was comparable between cross-sectional studies ($k = 7$, Hedges' $g = -0.35$, 95% CI [−0.52, −0.19], $p < .001$) and cross-sequential studies ($k = 6$, Hedges' $g = -0.35$, 95% CI [−0.59, −0.11], $p = .005$). There was no difference in effect sizes between studies using cross-sectional designs compared with studies using cross-sequential designs, $Q(1) = 0.003$, $p = .953$. We used meta-regression to evaluate the relationship between the five continuous moderators and effect size (see Table 5). As the rate of memory decay is known to be nonlinear (Rubin & Wenzel, 1996), delay interval and retention interval were log-transformed prior to analyses. None of these characteristics were significantly associated with effect size (all *p*s > .10).

Age-Related Differences in Memory for Canonical Categories

We also assessed age-related differences in flashbulb memory scores for four canonical categories: source, ongoing activity, location, and others present during the reception event (see Table 6). There were insufficient studies to evaluate age-related differences in scores for the time of the reception event ($k = 2$). Overall effect sizes indicated no significant differences between older and younger adults for memory for source, ongoing activity, location, or others present during the reception event. Full results of these analyses, including forest plots presenting effect sizes for individual studies, are provided in the online supplemental materials.

Age-Related Differences in Consistency of Flashbulb Memories

Eight studies assessed age-related differences in the consistency of flashbulb memories. Three studies found that older adults' flashbulb memories were significantly less consistent than younger adults', with Hedges' g ranging from −1.05 to −0.15 (see Table 7; for a forest plot presenting effect sizes for individual studies, see

Table 5
Analyses of Continuous Moderators of Flashbulb Memory Scores

Moderator	*k*	Slope	95% CI Lower bound	95% CI Upper bound	*p*
Methodological quality	14	0.001	−0.096	0.098	.987
Delay interval	11	0.031	−0.104	0.166	.654
Retention interval	6	0.001	−0.000	0.002	.145
M age: Older adults	13	−0.004	−0.075	0.067	.914
M age: Younger adults	13	0.007	−0.019	0.033	.602

Note. Delay interval and retention interval were log-transformed prior to analyses.

Table 6
Random Weighted Effect Sizes From Meta-Analyses of Flashbulb Memory Scores for Canonical Categories

Outcome	k	Hedges' g	95% CI Lower bound	95% CI Upper bound	Q (df)	I^2
Source						
Total (all studies)	5	0.07	−0.21	0.35	10.85 (4)*	63
Total (outliers excluded)	3	0.11	−0.09	0.31	0.60 (2)	0
Ongoing activity						
Total (all studies)	6	−0.12	−0.37	0.13	13.66 (5)*	63
Total (outlier excluded)	5	−0.03	−0.23	0.18	5.81 (4)	31
Location						
Total (all studies)	4	0.07	−0.25	0.39	8.43 (3)*	64
Total (outlier excluded)	3	0.19†	−0.02	0.40	1.65 (2)	0
Total (trim-and-fill procedure)		0.07	−0.14	0.28	5.73	
Others present						
Total (all studies)	3	−0.08	−0.53	0.37	7.65 (2)*	74
Total (trim-and-fill procedure)		−0.25	−0.72	0.21	14.11	

† $p < .10$. * $p < .05$.

online supplemental materials). The overall effect size indicated that older adults' memories were significantly less consistent than younger adults', with an effect size in the small-to-moderate range (Hedges' $g = -0.41$, 95% CI [−0.65, −0.17], $p = .001$). Outlier analyses using SAMD values and visual examination of the scree plot identified one study with an SAMD value greater than 2.24 which was also discrepant with the overall distribution (Cohen et al., 1994). After removal of this outlier, the overall effect size remained significant (Hedges' $g = -0.29$, 95% CI [−0.47, −0.11], $p = .002$). The Q statistic did not indicate significant heterogeneity among effect sizes, and the I^2 value of 0 suggested a very low degree of heterogeneity, so moderator analyses were not conducted for this outcome.

There was no evidence of publication bias in the included studies. The funnel plot appeared symmetrical, and the trim-and-fill procedure suggested no missing studies. However, the fail-safe N value of 11 was below the tolerance value (45), so this age-related difference should be interpreted with caution.

Age-Related Differences in Consistency of Memory for Canonical Categories

We also assessed age-related differences in the consistency of participants' memories for five canonical categories: source, ongoing activity, location, others present, and time of the reception event (time could be included here but not in the memory for canonical categories analysis because some studies reported consistency scores but not the Time 1 and Time 2 scores those were calculated from). Overall effect sizes indicated no significant differences in the consistency of older and younger adults' memories for source, location, or time (see Table 8). Results from four studies suggested that older adults' memories for the ongoing

Table 7
Random Weighted Effect Sizes From Meta-Analysis of Consistency of Flashbulb Memories

Study	Hedges' g	95% CI Lower bound	95% CI Upper bound	p	SAMD
Cohen et al. (1994)	−1.05***	−1.43	−0.67	<.001	−3.93[a]
Davidson et al. (2006)	−0.15	−0.68	0.37	.567	1.11
Davidson and Glisky (2002), Study 2	−0.20	−0.87	0.48	.574	0.63
Gerdy et al. (2007)	−0.35*	−0.69	−0.01	.046	0.38
Greene et al. (2018)	−0.23	−0.76	0.29	.387	1.01
Kensinger et al. (2006)	−0.57*	−1.14	0.00	.049	−0.67
Kvavilashvili et al. (2009), Study 3	−0.20	−0.54	0.14	.253	1.37
Otani et al. (2005)	−0.39	−1.09	0.31	.278	0.09

	k	Hedges' g	95% CI Lower bound	95% CI Upper bound	p	Q (df)	I^2
Total (all studies)	8	−0.41**	−0.65	−0.17	.001	14.42 (7)*	51
Total (outlier excluded)	7	−0.29**	−0.47	−0.11	.002	1.78 (6)	0

Note. SAMD = sample-adjusted meta-analytic deviancy statistic.
[a] Outlier excluded.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 8
Random Weighted Effect Sizes From Meta-Analyses of Consistency of Flashbulb Memories for Canonical Categories

Outcome	k	Hedges' g	95% CI Lower bound	95% CI Upper bound	Q (df)	I^2
Source						
Total (all studies)	4	−0.13	−0.73	0.48	20.71 (3)***	86
Total (outlier excluded)	3	0.13	−0.14	0.39	1.86 (2)	0
Total (trim-and-fill procedure)		0.06	−0.22	0.34	3.47	
Ongoing activity						
Total (all studies)	4	−0.61***	−0.95	−0.27	6.46 (3)†	54
Total (outlier excluded)	3	−0.43**	−0.70	−0.16	0.59 (2)	0
Total (trim-and-fill procedure)		−0.40**	−0.65	−0.15	1.07	
Location						
Total (all studies)	4	−0.27	−0.80	0.26	16.14 (3)**	81
Total (trim-and-fill procedure)		−0.07	−0.63	0.48	32.93	
Others present						
Total (all studies)	3	−0.18	−0.64	0.29	4.83 (2)†	59
Total (trim-and-fill procedure)		−0.51*	−1.00	−0.02	16.17	
Time						
Total (all studies)	3	−0.26	−0.61	0.10	3.00 (2)	33

† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

activity at the time of the reception event were significantly less consistent than younger adults'; this difference remained significant after exclusion of one outlier and adjustment for publication bias, with an effect size in the small-to-moderate range (Hedges' $g = -0.40$, 95% CI [−0.65, −0.15], $p = .002$). However, the fail-safe N value of 5 was below the tolerance value (25), so this difference should be interpreted with caution. Finally, although the overall effect size from three studies did not indicate a significant age-related difference in the consistency of memories for others present, after adjusting for publication bias, there was evidence that older adults' memories for others present were significantly less consistent than younger adults' (Hedges' $g = -0.51$, 95% CI [−1.00, −0.02], $p = .041$). Full results of these analyses and forest plots presenting effect sizes for individual studies are provided in the online supplemental materials.

Age-Related Differences in Encoding and Rehearsal Variables

Our final set of meta-analyses assessed age-related differences in five encoding and rehearsal variables associated with flashbulb memory formation and retention: investment in the flashbulb event, importance of the flashbulb event, the degree to which participants perceived the flashbulb event to be surprising, the perceived emotionality of the flashbulb event, and the degree of overt rehearsal (e.g., discussing the reception event with others). Overall effect sizes indicated no significant differences between older and younger adults for any of these variables (see Table 9). Full results of these analyses and forest plots presenting effect sizes for individual studies are provided in the online supplemental materials.

Discussion

This meta-analysis is the first to examine age-related differences in flashbulb memories. We found evidence for a small impairment in the presence of flashbulb memory for older adults compared with younger adults, which remained after removal of an outlier. Furthermore, older adults' flashbulb memories were significantly less consistent over time than those of younger adults. The latter finding, however, should be interpreted with caution given the number of included studies failed to exceed the tolerance value, which suggests that it is plausible that the true effect in the population may not be significant and that our finding might instead reflect omission of unpublished studies with null findings.

In these meta-analyses, we found that the effect size for the age-related differences in flashbulb memory scores and consistency is small (absolute values 0.30 and 0.29, respectively). Notably, these effect sizes are smaller than those that have been observed for age-related differences in predominantly laboratory-based memory studies; for example, working memory (0.80), prose recall (0.67), paired-associate and list recall (greater than 0.90, effect sizes all from Verhaeghen et al., 1993), associative memory (e.g., source, temporal order; 0.66 to 0.99, Old & Naveh-Benjamin, 2008), and prospective memory (0.54 to 0.72; Kliegel et al., 2008). The larger effect size for age-related impairments in other recent meta-analyses compared with the present one underscores the idea that carefully controlled laboratory tasks may suggest larger age-related deficits than older adults demonstrate on naturalistic memory tasks (Verhaeghen et al., 1993; West, 1986). To be sure, carefully controlled memory tasks are critical for moving forward our understanding of the basic mechanisms of cognitive aging. They must be supplemented, however, with studies of naturalistic memory tasks if we are to move forward our understanding of typical daily memory functioning of older adults. The present study is an example of the latter and suggests small age-related impairment in reporting the circumstances of learning significant news.

There are multiple reasons that age-related impairment in flashbulb memory may be smaller than those found with laboratory tasks designed to avoid ceiling effects in young adult

Table 9
Random Weighted Effect Sizes From Meta-Analyses of Encoding and Rehearsal Variables

Outcome	k	Hedges' g	95% CI Lower bound	95% CI Upper bound	Q (df)	I^2
Investment						
Total (all studies)	4	0.18	−0.34	0.69	15.79 (3)**	81
Importance						
Total (all studies)	11	−0.05	−0.18	0.07	10.86 (10)	8
Total (trim-and-fill procedure)		−0.12†	−0.26	0.03	22.32	
Surprise						
Total (all studies)	12	−0.01	−0.18	0.16	23.75 (11)*	54
Total (outliers excluded)	10	−0.09	−0.22	0.04	10.55 (9)	15
Total (trim-and-fill procedure)		−0.07	−0.22	0.08	15.03	
Emotionality						
Total (all studies)	13	0.16†	−0.01	0.32	25.26 (12)*	53
Total (outliers excluded)	10	0.11†	−0.01	0.23	6.48 (9)	0
Total (trim-and-fill procedure)		0.06	−0.05	0.17	11.12	
Overt rehearsal						
Total (all studies)	12	0.04	−0.14	0.21	19.99 (11)*	45
Total (outlier excluded)	11	0.09	−0.04	0.22	7.57 (10)	0
Total (trim-and-fill procedure)		0.00	−0.12	0.12	15.77	

† $p < .10$. * $p < .05$. ** $p < .01$.

groups. One is that age-related impairment in memory performance can be reduced when memory tasks include naturalistic materials so participants can rely on prior schemas (e.g., Castel, 2005). A second is that answering questions about learning significant news is relatively conversational so may not inspire stereotype threat with accompanying deficits that laboratory-based do, particularly those with intentional recall instructions that highlight memory is being assessed (e.g., Brubaker & Naveh-Benjamin, 2018; for a meta-analysis, see Armstrong, Gallant, Li, Patel, & Wong, 2017). Third, if one considers storytelling a cultural environmental support, the smaller age-related difference in flashbulb memory compared with the size of age-related differences found on laboratory-based memory tasks could be due to the reduced cognitive processing demands (e.g., Craik, 1994) required by sharing one's experience compared with recalling unrelated information. Fourth, flashbulb memories are typically about events that inspire emotional responses (see Table 1). Experiences or items that are emotionally charged are remembered better than neutral information, a phenomenon known as *emotional memory enhancement* (Kensinger, Allard, & Krendl, 2014). Emotional memory enhancement appears to be preserved into older adulthood (Charles, Mather, & Carstensen, 2003; Denburg, Buchanan, Tranel, & Adolphs, 2003; for a review, see Kensinger et al., 2014). Each of these possibilities may contribute to the reduced effect size in this meta-analysis of naturalistic memory phenomenon compared with the effect sizes in meta-analyses of predominantly laboratory-based memory (e.g., Verhaeghen et al., 1993). Although reduced, there was an age-related impairment in flashbulb memories.

Despite the presence of age-related differences in flashbulb memories, we did not find evidence for significant age-related differences in memory for most canonical categories. The only canonical category which yielded a significant age-related difference was the consistency of memory for ongoing activity at the time of the reception event. Moreover, these findings should be interpreted with caution, as the majority of these analyses indicated the presence of potential publication bias. This contrasting pattern between overall flashbulb memory scores and canonical categories scores may seem confusing, given that the flashbulb memory scores were typically comprised of scores on the canonical categories. Together, the data patterns suggest a case of the whole being more than the sum of the parts, that is, small age-related differences in canonical categories did not reveal a systematic pattern but the sum of them did. Another lens with which to view this contrast is that the overall age-related impairment in flashbulb memories does not seem localized in one or more specific canonical categories.

The lack of an age-related difference in source memory is especially interesting and can be interpreted in several ways. One is that older adults may rely on source guessing more than young adults do (e.g., Kuhlmann, Bayen, Meuser, & Kornadt, 2016), and life routines may support accurate guessing (e.g., I listen to the news every day). Another interpretation is that the lack of age-related differences in source memory is consistent with reports that the typical age gap in source memory can be overcome when conceptual cues are added (e.g., Rahhal, May, & Hasher, 2002). This latter interpretation aligns with socioemotional selectivity theory, in that the heightened emotion of flashbulb events may erase typical age-related deficits because older adults engage in greater attentional and elaboration processes with emotionally salient events (Carstensen & Turk-Charles, 1994). We do not argue for one interpretation over the other because they are not mutually exclusive. Moreover, flashbulb memory studies are not systematic investigations of source memory (e.g., there is only one item per participant). We simply note the connections between a case of naturalistic source memory and laboratory studies of age-related patterns in source memory.

By contrast, the significant age-related difference in ongoing activity consistency aligns with reported age-related differences in laboratory studies of activity memory (e.g., Earles & Kersten, 1999). The present report that older adults were less consistent than younger adults in reporting the activity that was interrupted by news of the flashbulb event may be explained, in part, by age-related differences in routines. For example, five studies examined memories for learning about the 9/11 attacks on the United States. The attacks happened in the morning, so young adults may recall the disruption to their morning work-related routines, whereas older adults would not have the benefit of such a cue to support their consistent recall over time.

The present meta-analyses also assessed encoding and rehearsal variables associated with flashbulb memory formation and retention, specifically investment, importance, surprise, emotionality, and overt rehearsal. Surprisingly, given the significant differences in flashbulb memory scores and consistency, we found no evidence for age-related differences in any encoding or rehearsal variables. These findings should be interpreted with caution, as four of these analyses indicated the presence of potential publication bias. These initial meta-analyses suggest that the age-related differences in flashbulb memory formation and retention may not occur because older adults perceive the events as less important, surprising, or emotional, or because older adults engage in less overt rehearsal or are less invested in the events. It is important to note, however, that we did not directly assess the relationship between these encoding and rehearsal variables and flashbulb memories, nor did we test whether these relationships differed between older and younger adults. The current literature does not provide sufficient evidence to conduct meta-analyses investigating these important topics. Thus, whether the reported age-related difference in flashbulb memory scores is due to encoding or retrieval factors, or both, has yet to be determined. Future research directly comparing the processes that underlie flashbulb memories in older and younger adults is necessary to determine whether age-related differences in flashbulb memories truly cannot be attributed to differences in these encoding and rehearsal variables. That open question may be unsurprising given that the flashbulb memory literature has yet to identify factors that are necessary and sufficient to elicit flashbulb memories (Hirst & Phelps, 2016).

That same review (Hirst & Phelps, 2016) emphasizes the link between flashbulb memories and social identity. Importantly, a key function of reminiscing is to clarify one's identity (additional functions include teach/inform, problem solving, boredom reduction, among others; e.g., Webster, 2003). Young adults report reminiscing for identity functions (e.g., finding meaning in life) more than older adults do (Cappeliez, Lavallée, & O'Rourke, 2001; Harris, Rasmussen, & Berntsen, 2014, Study 4; but see Vranić, Jelić, & Tonković, 2018, for a different pattern with a relatively young sample of older adults). Older adults report reminiscing for other functions, such as to teach/inform (e.g., Cappeliez et al., 2001; Webster & Gould, 2007), or the broader category of generativity (Harris et al., 2014, Study 4), more than young adults do. While many flashbulb memory studies assess variables such as how important the event is to participants and how often they rehearsed the memory, to our knowledge, studies have not yet unpacked why people reminisce about their flashbulb memories. This may be a useful measure to add into future studies of flashbulb memory, particularly those on aging. It is possible that the age-related impairment in flashbulb memories reported here is due, in part, to younger adults reminiscing for different purposes than older adults do. Perhaps reminiscing for identity (younger adults) links the recalled experience to enduring personal goals more than reminiscing for teaching/informing (older adults) does; such enduring personal goals are instrumental in guiding later autobiographical memory retrieval (Conway & Pleydell-Pearce, 2000). In a twist on Neisser's (1982) suggestion that flashbulb memories occur where personal and historical memories align, perhaps flashbulb memories occur where personal identity and social identity align.

A limitation of theses meta-analyses is the relatively small number of studies eligible for inclusion, especially studies reporting outcomes for specific canonical categories of flashbulb memories and encoding/rehearsal variables associated with flashbulb memory formation and retention. Only 21 studies, reflecting findings from 16 independent samples, reported adequate data from which effect sizes could be estimated. For one study (Cohen et al., 1994) data for the calculation of effect sizes had to be extracted from a published figure. Moreover, even when studies provided sufficient data for the calculation of overall effect sizes, data that could have been used for secondary analyses was not consistently reported or available from authors. For example, we were able to evaluate the age-related difference in consistency of memory for time of the reception event using effect sizes from three studies. Although these studies must have assessed memory for time at the initial assessment in order to generate these consistency scores, this data was not reported or available for two studies, thus we were unable to evaluate age differences in memory for time of the reception event. On the flip side, this limitation is an opportunity to suggest strategies that will enhance future meta-analyses. For example, difficulties in obtaining data highlight the importance of retaining original data sets; when possible researchers should post their raw data in a public repository (e.g., open science framework). Additionally, meta-analyses are becoming increasingly common, and thus it is imperative that researchers report outcome values from which effect sizes can be extracted. Retaining and sharing original data can also facilitate this process. Future meta-analytic reviews may address questions that are not addressed in initial publications, but for which a study can provide valuable data. Providing data in a comprehensive format allows for flexibility in the way that findings from studies can be used in support of future work. Finally, we recommend that future studies examining age-related differences in flashbulb memories report the means and standard deviations on the flashbulb memory measure for each group, as well as each canonical category and encoding and rehearsal variable assessed.

A related point is that, although there was variability in effect sizes among the included studies, exploratory moderator analyses failed to identify any characteristics of studies or samples that were systematically associated with age-related differences in flashbulb memory formation. Our ability to assess potential moderators was limited by the characteristics of the studies included in these analyses. Many studies did not report adequate demographic information regarding their samples. As a result, we were not able to assess the relationship between participant gender, education level, or race/ethnicity and effect size. Consistent with the American Psychological Association (APA) Journal Article Reporting Standards, we encourage researchers to collect and report comprehensive information regarding major demographic characteristics of participants, such as age, sex/gender, race/ethnicity, and socioeconomic status, as well as characteristics relevant to flashbulb memory, such as level of education and the presence of cognitive impairment (Appelbaum et al., 2018).

To increase overall methodological quality of studies in this area, we encourage the use of naïve coders when possible (acknowledging that some participant responses make that impossible for variables such as age group when older adults refer to grandchildren). A methodological improvement that is under researchers' control is the delay interval. The literature suggests that flashbulb memories are subject to distortions over time (Neisser & Harsch, 1992; Talarico & Rubin, 2003, 2007). Therefore, studies with shorter delay intervals likely offer more accurate depictions of flashbulb memory formation because of the increased likelihood of memory reconstruction and wrong time-slice errors in studies with longer delay intervals. We strongly recommend that researchers investigating flashbulb memory have instruments ready and Human Subjects Institutional Review Board approval so that Time 1 questions can be asked as soon as possible after an event occurs. Future work may also benefit from focusing on how often people think about flashbulb memories as compared with talking about them, given that people report thinking about the past more than talking about it and that pattern did not differ across age or gender (Bluck & Alea, 2009).

Finally, the scope of the present meta-analysis was focused on the presence of flashbulb memories, memory for canonical categories, consistency of the prior two measures over time, and exploration of encoding (e.g., importance, surprise) and retrieval variables. We did not assess other valuable and interesting questions related to aging and flashbulb memories. For example, in our analyses, we focus on flashbulb memories for public events. However, flashbulb memories may also form from important or surprising personal events (Demiray & Freund, 2015). Although several studies included in these meta-analyses reported memories for comparison events, the nature of these events varied widely: They included events that participants remembered vividly (e.g., Denver, Lane, & Cherry, 2010), personal events that occurred around the same time as the flashbulb event (e.g., Davidson & Glisky, 2002), and less emotional public events such as the Super Bowl (Kensinger et al., 2006). In comparing flashbulb memories for distinct types of events, it would be valuable for future research to consider whether these differences in memory vary with age. Furthermore, we acknowledge that confidence in flashbulb memories is an important characteristic to study; however, we leave for future research questions about aging and metamemory of flashbulb memories (e.g., whether confidence in flashbulb memories declines with age, whether the observed discrepancy between accuracy and confidence in flashbulb memories that has been reported in the literature [e.g., Talarico & Rubin, 2003] varies across the life span).

In summary, the present meta-analyses provide evidence for an age-related impairment in flashbulb memory, a naturalistic memory phenomenon. The results suggest that the difference Brown and Kulik (1977) noted between flashbulb memories and actual photographs—that the former are incomplete records—is larger for older adults compared with younger adults. Future research will clarify why that is with strengthened methodology (e.g., shorter delay between event and memory assessment) and, perhaps, with new questions (e.g., what function reminiscing about flashbulb memories serves for younger and older adults, and whether it varies by culture; cf. Bluck, 2015).

References

References marked with an asterisk indicate studies included in the meta-analyses.

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods, 16,* 270–301. http://dx.doi.org/10.1177/1094428112470848

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board Task Force report. *American Psychologist, 73,* 3–25. http://dx.doi.org/10.1037/amp0000191

Armstrong, B., Gallant, S. N., Li, L., Patel, K., & Wong, B. I. (2017). Stereotype threat effects on older adults' episodic and working memory: A meta-analysis. *The Gerontologist, 57*(Suppl. 2), S193–S205. http://dx.doi.org/10.1093/geront/gnx056

Bluck, S. (2015). Going global: Functions of autobiographical remembering world tour. *Memory, 23,* 111–118. http://dx.doi.org/10.1080/09658211.2014.937721

Bluck, S., & Alea, N. (2009). Thinking and talking about the past: Why remember? *Applied Cognitive Psychology, 23,* 1089–1104. http://dx.doi.org/10.1002/acp.1612

*Bohn, A., & Berntsen, D. (2007). Pleasantness bias in flashbulb memories: Positive and negative flashbulb memories of the fall of the Berlin Wall among East and West Germans. *Memory & Cognition, 35,* 565–577. http://dx.doi.org/10.3758/BF03193295

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *The Journals of Gerontology Series B, Psychological Sciences and Social Sciences, 60,* P223–P233. http://dx.doi.org/10.1093/geronb/60.5.P223

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis* (1st ed.). Chichester, UK: Wiley. http://dx.doi.org/10.1002/9780470743386

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2017). *Comprehensive meta-analysis* (Version 3) [Computer software]. Englewood, NJ: Biostat. Retrieved from https://www.meta-analysis.com/?gclid=Cj0KCQjw09HzBRDrARIsAG60GP_81D4seyQ6KrhNfCxodM39kbKDPhPJ3-NcwzC_3TwNufWssj5lbycaAtsnEALw_wcB

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5,* 73–99. http://dx.doi.org/10.1016/0010-0277(77)90018-X

Brubaker, M. S., & Naveh-Benjamin, M. (2018). The effects of stereotype threat on the associative memory deficit of older adults. *Psychology and Aging, 33,* 17–29. http://dx.doi.org/10.1037/pag0000194

Cappeliez, P., Lavallée, R., & O'Rourke, N. (2001). Functions of reminiscence in later life as viewed by young and old adults. *Canadian Journal on Aging, 20,* 577–589. http://dx.doi.org/10.1017/S0714980800012320

Carstensen, L. L., & Turk-Charles, S. (1994). The salience of emotion across the adult life span. *Psychology and Aging, 9,* 259–264. http://dx.doi.org/10.1037/0882-7974.9.2.259

Castel, A. D. (2005). Memory for grocery prices in younger and older adults: The role of schematic support. *Psychology and Aging, 20,* 718–721. http://dx.doi.org/10.1037/0882-7974.20.4.718

Charles, S. T., Mather, M., & Carstensen, L. L. (2003). Aging and emotional memory: The forgettable nature of negative images for older adults. *Journal of Experimental Psychology: General, 132,* 310–324. http://dx.doi.org/10.1037/0096-3445.132.2.310

Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics, 10,* 101–129. http://dx.doi.org/10.2307/3001666

*Cohen, G., Conway, M. A., & Maylor, E. A. (1994). Flashbulb memories in older adults. *Psychology and Aging, 9,* 454–463. http://dx.doi.org/10.1037/0882-7974.9.3.454

Conway, A. R. A., Skitka, L. J., Hemmerich, J. A., & Kershaw, T. C. (2009). Flashbulb memory for 11 September 2001. *Applied Cognitive Psychology, 23,* 605–623. http://dx.doi.org/10.1002/acp.1497

Conway, M. A., Anderson, S. J., Larsen, S. F., Donnelly, C. M., McDaniel, M. A., McClelland, A. G. R., . . . Logie, R. H. (1994). The formation of flashbulb memories. *Memory & Cognition, 22,* 326–343. http://dx.doi.org/10.3758/BF03200860

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review, 107,* 261–288. http://dx.doi.org/10.1037/0033-295X.107.2.261

Craik, F. I. M. (1994). Memory changes in normal aging. *Current Directions in Psychological Science, 3,* 155–158. http://dx.doi.org/10.1111/1467-8721.ep10770653

*Davidson, P. S. R., Cook, S. P., & Glisky, E. L. (2006). Flashbulb memories for September 11th can be preserved in older adults. *Aging, Neuropsychology and Cognition, 13,* 196–206. http://dx.doi.org/10.1080/13825580490904192

*Davidson, P. S. R., & Glisky, E. L. (2002). Is flashbulb memory a special instance of source memory? Evidence from older adults. *Memory, 10,* 99–111. http://dx.doi.org/10.1080/09658210143000227

Demiray, B., & Freund, A. M. (2015). Michael Jackson, Bin Laden and I: Functions of positive and negative, public and private flashbulb memories. *Memory, 23,* 487–506. http://dx.doi.org/10.1080/09658211.2014.907428

Denburg, N. L., Buchanan, T. W., Tranel, D., & Adolphs, R. (2003). Evidence for preserved emotional memory in normal older persons. *Emotion, 3,* 239–253. http://dx.doi.org/10.1037/1528-3542.3.3.239

*Denver, J. Y., Lane, S. M., & Cherry, K. E. (2010). Recent versus remote: Flashbulb memory for 9/11 and self-selected events from the reminiscence bump. *International Journal of Aging & Human Development, 70,* 275–297. http://dx.doi.org/10.2190/AG.70.4.a

Duval, S., & Tweedie, R. (2000a). A nonparametric "Trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95,* 89–98. http://dx.doi.org/10.2307/2669529

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56,* 455–463. http://dx.doi.org/10.1111/j.0006-341X.2000.00455.x

Earles, J. L., & Kersten, A. W. (1999). Processing speed and adult age differences in activity memory. *Experimental Aging Research, 25,* 243–253. http://dx.doi.org/10.1080/036107399244011

Er, N. (2003). A new flashbulb memory model applied to the Marmara earthquake. *Applied Cognitive Psychology, 17,* 503–517. http://dx.doi.org/10.1002/acp.870

Finkenauer, C., Luminet, O., Gisle, L., El-Ahmadi, A., Van Der Linden, M., & Philippot, P. (1998). Flashbulb memories and the underlying mechanisms of their formation: Toward an emotional-integrative model. *Memory & Cognition, 26,* 516–531. http://dx.doi.org/10.3758/BF03201160

Fraundorf, S. H., Hourihan, K. L., Peters, R. A., & Benjamin, A. S. (2019). Aging and recognition memory: A meta-analysis. *Psychological Bulletin, 145,* 339–371. http://dx.doi.org/10.1037/bul0000185

*Gerdy, J. E., Multhaup, K. M., & Ivey, P. C. (2007). *Flashbulb memories in older and younger adults*. (Unpublished manuscript). Davidson College, Davidson, NC.

*Greene, C. M., Loftus, E., Grady, R. H., & Levine, L. (2018, November 15–18). *Flashbulb memories of a national referendum on abortion*. Paper presented at the 59th Annual Meeting of the Psychonomic Society, New Orleans, LA. Retrieved from https://cdn.ymaws.com/www.psychonomic.org/resource/resmgr/annual_meeting/2018_meeting/program/ps2018abstracts25.pdf

Harris, C. B., Rasmussen, A. S., & Berntsen, D. (2014). The functions of autobiographical memory: An integrative approach. *Memory, 22,* 559–581. http://dx.doi.org/10.1080/09658211.2013.806555

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators, *6,* 107–128. http://dx.doi.org/10.2307/1164588

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal (Clinical Research Ed.), 327,* 557–560. http://dx.doi.org/10.1136/bmj.327.7414.557

Hirst, W., & Phelps, E. A. (2016). Flashbulb memories. *Current Directions in Psychological Science, 25,* 36–41. http://dx.doi.org/10.1177/0963721415622487

Hirst, W., Phelps, E. A., Meksin, R., Vaidya, C. J., Johnson, M. K., Mitchell, K. J., . . . Olsson, A. (2015). A ten-year follow-up of a study of memory for the attack of September 11, 2001: Flashbulb memories and memories for flashbulb events. *Journal of Experimental Psychology: General, 144,* 604–623. http://dx.doi.org/10.1037/xge0000055

Huffcutt, A. I., & Arthur, W. (1995). Development of a new outlier statistic for meta-analytic data. *Journal of Applied Psychology, 80,* 327–334. http://dx.doi.org/10.1037/0021-9010.80.2.327

Kempe, M., Kalicinski, M., & Memmert, D. (2015). Naturalistic assessment of everyday memory performance among older adults. *Experimental Aging Research, 41,* 426–445. http://dx.doi.org/10.1080/0361073X.2015.1053766

Kensinger, E. A., Allard, E. R., & Krendl, A. C. (2014). The effects of age on memory for socioemotional material: An affective neuroscience perspective. In P. Verhaeghen & C. Hertzog (Eds.), *The Oxford handbook of emotion, social cognition, and problem solving in adulthood* (pp. 26–46). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199899463.013.011

*Kensinger, E. A., Krendl, A. C., & Corkin, S. (2006). Memories of an emotional and a nonemotional event: Effects of aging and delay interval. *Experimental Aging Research, 32,* 23–45. http://dx.doi.org/10.1080/01902140500325031

Kızılöz, B. K., & Tekcan, A. I. (2013). Canonical categories in flashbulb memories. *Applied Cognitive Psychology, 27,* 352–359. http://dx.doi.org/10.1002/acp.2913

Kliegel, M., Jäger, T., & Phillips, L. H. (2008). Adult age differences in event-based prospective memory: A meta-analysis on the role of focal versus nonfocal cues. *Psychology and Aging, 23,* 203–208. http://dx.doi.org/10.1037/0882-7974.23.1.203

Kuhlmann, B. G., Bayen, U. J., Meuser, K., & Kornadt, A. E. (2016). The impact of age stereotypes on source monitoring in younger and older adults. *Psychology and Aging, 31,* 875–889. http://dx.doi.org/10.1037/pag0000140

*Kvavilashvili, L., Mirani, J., Schlagman, S., Erskine, J. A. K., & Kornbrot, D. E. (2010). Effects of age on phenomenology and consistency of flashbulb memories of September 11 and a staged control event. *Psychology and Aging, 25,* 391–404. http://dx.doi.org/10.1037/a0017532

*Kvavilashvili, L., Mirani, J., Schlagman, S., Wellsted, D., & Kornbrot, D. E. (2009). *"I can remember this as if it happened yesterday": Comparing flashbulb memories of very recent and distance events*. (Unpublished manuscript). Hertfordshire, UK: Department of Psychology, University of Hertfordshire.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: SAGE.

Luminet, O. (2018). Using structural equation modelling approaches to better understand the formation of flashbulb memories. In O. Luminet & A. Curci (Eds.), *Flashbulb memories: New challenges and future perspectives* (pp. 48–72). New York, NY: Routledge.

Martschuk, N., & Sporer, S. L. (2018). Memory for faces in old age: A meta-analysis. *Psychology and Aging, 33,* 904–923. http://dx.doi.org/10.1037/pag0000282

McAlister, C., & Schmitter-Edgecombe, M. (2016). Cross-sectional and longitudinal analyses of everyday memory lapses in older adults. *Aging, Neuropsychology and Cognition, 23,* 591–608. http://dx.doi.org/10.1080/13825585.2015.1132669

Neisser, U. (1982). Snapshots or benchmarks? In U. Neisser (Ed.), *Memory observed: Remembering in natural contexts* (pp. 43–48). San Francisco, CA: Freeman.

Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about Challenger. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of "flashbulb" memories* (pp. 9–31). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511664069.003

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging, 23,* 104–118. http://dx.doi.org/10.1037/0882-7974.23.1.104

Ossher, L., Flegal, K. E., & Lustig, C. (2013). Everyday memory errors in older adults. *Aging, Neuropsychology and Cognition, 20,* 220–242. http://dx.doi.org/10.1080/13825585.2012.690365

*Otani, H., Kusumi, T., Kato, K., Matsuda, K., Kern, R. P., Widner, R., Jr., & Ohta, N. (2005). Remembering a nuclear accident in Japan: Did it trigger flashbulb memories? *Memory, 13,* 6–20. http://dx.doi.org/10.1080/09658210344000495

Rahhal, T. A., May, C. P., & Hasher, L. (2002). Truth and character: Sources that older adults can remember. *Psychological Science, 13,* 101–105. http://dx.doi.org/10.1111/1467-9280.00419

Rhodes, S., Greene, N. R., & Naveh-Benjamin, M. (2019). Age-related differences in recall and recognition: A meta-analysis. *Psychonomic Bulletin & Review, 26,* 1529–1547. http://dx.doi.org/10.3758/s13423-019-01649-y

Rohatgi, A. (2019). WebPlotDigitizer (Version 4.2) [Computer software]. Retrieved from https://automeris.io/WebPlotDigitizer

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86,* 638–641. http://dx.doi.org/10.1037/0033-2909.86.3.638

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review, 103,* 734–760. http://dx.doi.org/10.1037/0033-295X.103.4.734

Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging, 10,* 527–539. http://dx.doi.org/10.1037/0882-7974.10.4.527

Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., . . . Higgins, J. P. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal (Clinical Research Ed.), 343,* d4002. http://dx.doi.org/10.1136/bmj.d4002

Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science, 14,* 455–461. http://dx.doi.org/10.1111/1467-9280.02453

Talarico, J. M., & Rubin, D. C. (2007). Flashbulb memories are special after all; in phenomenology, not accuracy. *Applied Cognitive Psychology, 21,* 557–578. http://dx.doi.org/10.1002/acp.1293

*Tekcan, A. I. (in press). *Age differences in flashbulb memory formation for seven public events* (Unpublished raw data). Boğaziçi University, Istanbul, Turkey.

*Tekcan, A. İ., & Peynircioğlu, Z. F. (2002). Effects of age on flashbulb memories. *Psychology and Aging, 17,* 416–422. http://dx.doi.org/10.1037/0882-7974.17.3.416

Uttl, B. (2011). Transparent meta-analysis: Does aging spare prospective memory with focal vs. non-focal cues? *PLoS ONE, 6*(2), e16618. http://dx.doi.org/10.1371/journal.pone.0016618

Verhaeghen, P., Marcoen, A., & Goossens, L. (1993). Facts and fiction about memory aging: A quantitative integration of research findings. *Journal of Gerontology, 48,* 157–171. http://dx.doi.org/10.1093/geronj/48.4.P157

Vranić, A., Jelić, M., & Tonković, M. (2018). Functions of autobiographical memory in younger and older adults. *Frontiers in Psychology, 9,* 219. http://dx.doi.org/10.3389/fpsyg.2018.00219

Webster, J. D. (2003). The reminiscence circumplex and autobiographical memory functions. *Memory, 11,* 203–215. http://dx.doi.org/10.1080/741938202

Webster, J. D., & Gould, O. (2007). Reminiscence and vivid personal memories across adulthood. *International Journal of Aging & Human Development, 64,* 149–170. http://dx.doi.org/10.2190/Q8V4-X5H0-6457-5442

West, R. L. (1986). Everyday memory and aging. *Developmental Neuropsychology, 2,* 323–344. http://dx.doi.org/10.1080/87565648609540352

*Wolters, G., & Goudsmit, J. J. (2005). Flashbulb and event memory of September 11, 2001: Consistency, confidence and age effects. *Psychological Reports, 96,* 605–619. http://dx.doi.org/10.2466/pr0.96.3.605-619

Wright, D. B., Gaskell, G. D., & O'Muircheartaigh, C. A. (1998). Flashbulb memory assumptions: Using national surveys to explore cognitive phenomena. *British Journal of Psychology, 89,* 103–121. http://dx.doi.org/10.1111/j.2044-8295.1998.tb02675.x

*Yarmey, A. D., & Bull, M. P., III. (1978). Where were you when President Kennedy was assassinated? *Bulletin of the Psychonomic Society, 11,* 133–135. http://dx.doi.org/10.3758/BF03336788

Received July 2, 2019
Revision received February 27, 2020
Accepted March 15, 2020 ■