Ψ Psychology Press
Taylor & Francis Group

Reminiscence bump in memory for public events

Steve M. J. Janssen and Jaap M. J. Murre
University of Amsterdam, Amsterdam, The Netherlands

Martijn Meeter
Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

People tend to recall more personal events from adolescence and early adulthood than from other lifetime periods. Most evidence suggests that differential encoding causes this reminiscence bump. However, the question why personal events are encoded better in those periods is still unanswered. To shed more light on this discussion, we examined memory for public events. Since it is often impossible to ascertain that queried events are equally difficult, we circumvented the issue of equivalence by calculating deviation scores for each trial. We found that participants more frequently answered questions correctly about events that occurred in the period in which they were between 10 and 25 years old. Furthermore, we found that the reminiscence bump was more pronounced for cued recall than for recognition. We argue that these results support the biological account that events are stored better, because the memory system is working more efficiently during adolescence and early adulthood. These results do not falsify the other accounts for differential encoding, because they are not mutually exclusive.

People speak of autobiographical memory when they are referring to the memories they have of their own life experiences (Robinson, 1986). Autobiographical memory does not only consist of personal memories that are remembered vividly, but also of autobiographical facts (Brewer, 1986). Some researchers have examined the contents of autobiographical memories (e.g., Fitzgerald, 1988; Niedźwieńska, 2003; Robinson, 1976), whereas other researchers have focused on the temporal distribution of memories of personal events across the lifespan (e.g., Janssen, Chessa, &

Murre, 2005; Rubin, Rahhal, & Poon, 1998; Rubin, Wetzler, & Nebes, 1986; Rybash, 1999).

Such lifetime distributions tend to have three characteristics. The first characteristic is childhood or infantile amnesia. People recall very few personal events from the first 3 or 4 years of their life. The second characteristic is referred to as the reminiscence bump. People tend to recall many personal events from adolescence and early adulthood. They usually recall more memories from these lifetime periods than from later periods. However, people also recall many personal events from recent years. This characteristic is called the recency effect. Unsurprisingly, the more recent an event is, the more likely it is to be remembered.

There has been some debate about whether recent and remote autobiographical memories are to the same extent episodic. Episodic memories are generally seen as personal events bound to a spatial and temporal context, while semantic memories or autobiographical facts contain knowledge about personal events (Brewer, 1986; Tulving, 1972). Cermak (1984) considered recent personal events to be episodic, while he thought that most remote memories were semantic. Schooler, Shiffrin, and Raaijmakers (2001) proposed a theory about how episodic memories could lose their contextual information and so become semantic memories. Many theorists assume that episodic memories are progressively modified in neocortical regions until they are independent of the hippocampal complex (e.g., Meeter & Murre, 2004; Murre, 1996, 1997; Rosenbaum, Winocur, & Moscovitch, 2001).

However, Rybash and Monaghan (1999) presented 40 participants, who were between the ages 70 and 75, with 18 cue words. For each cue word, the participants had to describe the memory that came to mind first. Then, they had to indicate whether they remembered the event or knew that event had occurred. Finally, they had to date the events. Rybash and Monaghan found that the distribution of remembered, episodic memories as well as the distribution of known semantic memories had a reminiscence bump and a recency effect. The reminiscence bump did not only consist of semantic knowledge, while the recency effect did not only consist of episodic memories.

CAUSES OF THE REMINISCENCE BUMP

There are at least three plausible explanations for the occurrence of the reminiscence bump. The first explanation is that in certain calendar years more memorable events happened than in others. There is indeed some evidence for external influences on the temporal distribution of personal events. Schrauf and Rubin (2001) analysed the distribution of personal events of Hispanic participants who immigrated to the United States as an

adult. Most memories came from around the period in which they immigrated. Rubin and Berntsen (2003) asked Danish participants in which event they felt the most scared, proud, jealous, loved, and angry in their life. They found that participants, who were 70 years or older, recalled the most scariest and traumatic events from the period when Denmark was occupied by the Germans during the Second World War. Conway and Haque (1999) asked Bangladeshi participants to describe the personal event that first comes to mind when presented with a cue word. They found a bump that coincided with the Bengali independence war. However, independent of this bump, they also found the classic reminiscence bump, since most memories came from early adulthood. This is just one example of the robustness of the reminiscence bump. It can be found in the distributions of different age groups in one study, even though for these groups the reminiscence bump periods fall in different calendar years (Rubin et al., 1986). This strongly suggests that the calendar year explanation is insufficient, although certain public events may affect the temporal distribution of personal events.

The second explanation is referred to as resampling. It states that at a certain age people start reminiscing about the period that they were adolescents or young adults (Rubin et al., 1986). However, this explanation has some shortcomings. First, it is unclear why people mainly reminisce about events from adolescence and early adulthood and not about events from other lifetime periods. Furthermore, Hyland and Ackerman (1988), Merriam and Cross (1982), and Webster and McCall (1999) showed that young adults spend an equal amount of time on reminiscing as older adults. Finally, the resampling explanation is unlikely to be the sole explanation of the reminiscence bump, because the distributions of adolescents and young adults display reminiscence bumps when one removes the recency effect from the lifetime distributions (Janssen et al., 2005).

The third explanation is called differential encoding. It states that events in adolescence and early adulthood are stored better than in other lifetime periods (Rubin et al., 1986, 1998). Four mechanisms can be given for differential encoding, but these mechanisms are not mutually exclusive. First, there are more novel events in those lifetime periods (cognitive account), such as the first driving lesson or the first kiss. These first-time experiences are encoded more strongly because they can be used later in life as exemplars when people experience similar events (Pillemer, 2001; Robinson, 1992). Second, people form their identity during those lifetime periods (identity formation account). Conway (2005; also see Fitzgerald, 1988) hypothesises that many self-defining memories, which are vivid and emotional memories of personal events that have a large impact on the identity of a person (Conway, Singer, & Tagini, 2004), come from adolescence and early adulthood. Third, more transitional events occur during those lifetime periods (life scripts account). When people are asked to

date their most important personal events, they recall events, such as graduation, wedding, and retirement (Berntsen & Rubin, 2002; Rubin & Berntsen, 2003). These important events are often positive, but sometimes negative. The positive events usually occur in early adulthood, whereas the negative events can occur at any moment in people's lives (Berntsen & Rubin, 2004; Rubin & Berntsen, 2003). Fourth, the events are stored better, because the brain works best in those lifetime periods (biological account), causing the memory system to work more efficiently.

Differential encoding and resampling

Our view, however, is that both differential encoding *and* resampling influence the reminiscence bump (Janssen et al., 2005; Janssen, Chessa, & Murre, in press; Janssen & Murre, 2007). We hypothesised that events are encoded more strongly during adolescence and early adulthood, because the memory system is working more efficiently during those lifetime periods (i.e., biological account). Furthermore, we hypothesised that events from adolescence and early adulthood are resampled more frequently than events from other lifetime periods, because events that have been encoded strongly have a larger likelihood to be retrieved (e.g., Anderson & Schooler, 1991) or because these events are self-defining moments (i.e., identity formation account) or transitional events (i.e., life scripts account).

In Janssen et al. (2005), we presented 10 cue words to participants of different ages. We removed the recency effect from the lifetime distributions, because the recency effect coincides with the reminiscence bump for adolescents and young adults. We found reminiscence bumps in the distributions of all age groups, including those adolescents and young adults. However, the size of the reminiscence bump increased as participants became older.

In Janssen and Murre (2007), we found that participants recalled relatively fewer remarkable events (i.e., novel, emotional, positive, and important events) than mundane events (i.e., regular, unemotional, neutral, and unimportant events) from adolescence. We also found that the difference between remarkable and mundane events becomes smaller as participants became older, because mundane events were forgotten faster than remarkable events.

The reminiscence bump has not only been found in the distribution of personal events, but also in the distribution of favourite books, movies, and records (Holbrook & Schindler, 1989, 1996; Janssen et al., in press; Larsen, 1996; Rubin et al., 1998; Sehulster, 1996; Smith, 1994). In Janssen et al. (in press), we asked participants to name their three favourite books, movies, and records and to indicate when they first encountered each item. We found

that the distributions of the favourite books, movies, and records peaked in the period in which the participants were between 16 and 20 years old. Interestingly, we found a larger recency effect for books than for records and a larger reminiscence bump for records than for books. Because people read their favourite books only two or three times, books are hardly rehearsed and therefore forgotten easily. People tend to listen to their favourite records numerous times, causing records to be rehearsed frequently and therefore to be more resistant to forgetting.

Because the brain works best during adolescence and early adulthood, the distributions of autobiographical memory of adolescents and young adults have reminiscence bumps as well (Janssen et al., 2005), mundane events are overrepresented in adolescence (Janssen & Murre, 2007), and there are reminiscence bumps in the distributions of favourite books, movies, and records (Holbrook & Schindler, 1989, 1996; Janssen et al., in press; Larsen, 1996; Rubin et al., 1998; Sehulster, 1996; Smith, 1994). Events from adolescence and early adulthood are sampled more frequently, because they have a larger likelihood to be recalled (e.g., Anderson & Schooler, 1991), or because they are self-defining memories (Conway, 2005; Fitzgerald, 1988) or transitional events (Berntsen & Rubin, 2002; Rubin & Berntsen, 2003). Because the events are recalled more often, they become even stronger as time passes, which explains the increase in the reminiscence bump in the distribution of autobiographical memory of older adults (Janssen et al., 2005), the relative increase of remarkable events from adolescence as people become older (Janssen & Murre, 2007), and the larger reminiscence bump in the distribution of favourite records than the reminiscence bump in the distribution of favourite books (Janssen et al., in press).

In this study, we will investigate the hypothesis that events are encoded more strongly during adolescence and early adulthood, because the memory system is working more efficiently during those lifetime periods. If so, then the reminiscence bump should also be found in distributions of nonauto-biographical memory domains, such as memory for public events. The other accounts for differential encoding do not specifically predict a reminiscence bump in the distribution of memory for public events, because they refer to first-time experiences, self-defining moments, or transitional events. We assume that memory for public events, like autobiographical memory (Brewer, 1986; Cermak, 1984), is a part of episodic as well as semantic memory. People can remember some public events vividly, including the specific temporal-spatial context in which the information was acquired (e.g., Brown & Kulik, 1977; Neisser, 1982); of other events they can only recall facts, without recalling the specific temporal-spatial context in which the information was acquired.

MEMORY FOR PUBLIC EVENTS

Earlier research into age effects in the temporal distribution of memory for public events has yielded consistent results. Howes and Katz (1988), Squire (1974), and Warrington and Sanders (1971) all found similar age and ageing effects using questionnaires of news events. Adults performed better on remote questions than adolescents and young adults, but they performed relatively worse on very recent questions. As adult participants became older, they performed poorer on very recent questions.

These experiments did not directly address the issue whether people encode information about public events better in the period in which they were adolescents or young adults than in subsequent lifetime periods, but Holmes and Conway (1999) and Schuman, Belli, and Bischoping (1997) did address this issue directly. Schuman et al. presented 11 events that were represented by names or persons to various representative populations of 1000–2000 participants. The participants were asked if they knew the event that the name or person represented and, if so, to describe the event. Schuman et al. examined whether participants gave more accurate descriptions of the events, when the event occurred in the period in which they were between 15 and 30 years old. They found that this assumption held up for three events, but not for the other eight events, because those other events were very recent or very remote. Holmes and Conway examined memory for public events in two experiments. In the first experiment, participants had to recall as many personal and public events as possible, similarly to a fluency test. Afterwards the participants had to date these personal and public events. The results showed that participants recalled most personal events from the period in which they were 20–30 years old and most public events from the period in which they were 10–20 years old. In the second experiment, participants had to complete ambiguous names. The completed names often referred to somebody who was famous in the period in which the participants were 10–20 years old.

Longmore, Knight, and Longmore (1990) identified three potential problems, which make it difficult to draw conclusions from the above-mentioned results. First, the events must be time-specific. Second, the events must have been learned at the time of the event. This criterion excludes very important events, which are, almost by definition, also learned by later generations. Third, the events from different periods must be learned equally well (i.e., equal item difficulty). If this is not certain, one never knows whether, for example, people recall recent events better than remote events because of a difference in original learning rate or because of a difference in age. To eliminate the problem of unequal item difficulty, Squire and Slater (1975) chose to ask questions about similar types of events from each time period (i.e., categorical question method). They asked standardised

questions about television shows and racehorses to high school children, university students, and adults. They found that high school children recognised more names of recent television shows than adults, whereas adults recognised more names of remote television shows than high school children. University students recognised as many names of recent racehorses as adults did, but adults recognised more names of remote racehorses than university students did.

Rubin et al. (1998); also see Rubin, Schulkind, & Rahhal, 1999) followed Squire and Slater's (1975) logic and made five categories of 60 standardised multiple-choice questions (e.g., "Which baseball teams played in the World Series in year X?" or "Which picture won the Academy Award for best picture in year X?"). They tested the questions with four groups. The first two groups consisted of 20-year-olds and 70-year-olds tested in 1984, while the last two groups consisted of different 20-year-olds and 70-year-olds tested in 1994. Every group was given all 300 questions, even those questions referring to events that occurred before their birth. Rubin et al. found a reminiscence bump in the temporal distribution of the proportion correct. Participants were most accurate when the questions came from their early adulthood.

Asking for events from one specific category, such as the names of racehorses, Oscar winners, or baseball teams in the World Series, makes the equality of item difficulty a priori more likely, but there are nevertheless two potential pitfalls to this method. The first problem is referred to as intergenerational interest. Intergenerational comparisons by necessity rely on a quasiexperimental setup. Interest in a certain category of events can wax or wane, which may introduce bias. If baseball and the Oscars were less widely reported on in one period than in another, one could find spurious differences in responding for different generations. For example, the World Series between the New York Mets and the New York Yankees (i.e., Subway Series) may attract more media attention than other World Series. The second problem is referred to as intergeneration matching. It can never be excluded that some difference in the matched populations accounts for differences in scores. This is most obvious with gender differences that are found in the recognition of sports events (Howes & Katz, 1988; Rubin et al., 1999). If one wants to compare the performance on sports questions of different age groups, one has to make sure that both age groups consist of the same proportion males and females, as Howes and Katz did, or one has to analyse the results of men and women separately, as Rubin et al. did. However, gender is just one of many demographic variables that could influence the results. For example, Schuman et al. (1997) found that African Americans gave more detailed descriptions about the event involving Rosa Parks than Caucasian Americans did. The reminiscence bump in the

temporal distribution of memory for public events has therefore not been established unquestionably.

YEARLY NEWS MEMORY TEST

In this study, we will use a conditional probabilities approach to resolve the potential problem of unequal item difficulty rather than relying on what one could call the categorical question method. A very large group of participants from different age groups has been given a semirandom selection of 30 open-ended and multiple-choice questions about a wide array of public events from the last 56 years (i.e., 1950–2006). We have chosen to ask open-ended as well as multiple-choice questions, because the difference between younger and older participants is larger on episodic recall tests than on episodic recognition tests (e.g., Craik & McDowd, 1987; Nyberg et al., 2003; Spaan, Raaijmakers, & Jonker, 2003).

We will calculate the proportion correct for all trials, for each participant and for each question. Using these three proportions correct, we will estimate an expectancy probability for each trial. These expectancy probabilities will be compared with the observed scores on each trial, resulting in deviation scores. These deviation scores will then be used in the statistical analyses. This method circumvents problems of unequal item difficulty, intergenerational interest, and matching, because the mean score of each question and each participant is estimated and used to correct the expectancy probability.

Longmore et al. (1990) gave three criteria for research into the temporal distribution of memory for public events. The events have to be time-specific, learned at the time of the event, and learned equally well. Unfortunately, most public events have consequences, such as court cases, resignation of responsible people, or unveiling of monument commemorating the event, making the public events less time-specific. People are also reminded of events when similar events occur, which can cause additional learning of the event. For example, when the shooting at Virginia Tech happened, many newspaper articles and news broadcasts recalled the shootings at Columbine High School. The news events can also be rehearsed through books or documentaries. We assume that these effects consist mainly of consolidating existing memories, because the aftermath and reoccurrences have less impact and news coverage than the initial news event. They strengthen existing memories rather than form new memories. We feel that these effects are small, but we should keep in mind that if this is not the case, they lead to a flattening of the reminiscence bump in the distribution.

For each public event, participants can be either younger than 10 years, between 10 and 25 years old, or older than 25 years when the event occurred.

We expect to find that questions that were presented to participants who were older than 10 years at the occurrence of the event were more frequently answered correctly than questions that were presented to participants who were younger than 10 years. If the biological account of differential encoding is correct, we would furthermore find that questions that were presented to participants who were between 10 and 25 years old were more frequently answered correctly than questions that were presented to participants who were older than 25 years. Finally, we expect to find that the reminiscence bump is stronger in the results of the open-ended questions than in the results of the multiple-choice questions.

METHOD

Participants

The dynamic public events questionnaire, which we called the *Yearly News Memory Test* (YNMT), was administered in Dutch via the Internet, where it is still available at http://memory.uva.nl/testpanel/ltmt/. Participants could come into contact with our website in at least four ways: (1) through links on other websites, (2) through search engines, (3) through promotion in traditional media, such as articles in newspapers and magazines, which included our web address, or (4) through word of mouth. At the end of the test, participants could invite relatives, friends, and colleagues by sending them standardised e-mails. Furthermore, we invited participants who had taken other tests on our website, such as the Daily News Memory Test (Meeter, Murre, & Janssen, 2005), the Galton-Crovitz test (Janssen et al., 2005, 2006; Janssen & Murre, 2007) or the Favourites Questionnaire (Janssen et al., in press), to take this test as well.

The results analysed in this paper were gathered between April 2005 and June 2006. During this period, 1334 Dutch participants between the ages 16 and 75 completed the test. The results of participants who lived abroad or were younger than 16 years or older than 75 years were not analysed, because those age groups consisted of too few participants. The average age of the participants was 42.9 years. We divided participants into six age groups (16–25 years, $N = 266$; 26–35 years, $N = 236$, 36–45 years, $N = 259$; 46–55 years, $N = 338$; 56–65 years, $N = 179$; 66–75 years, $N = 56$). More than half of the participants was female (60.1%). Participants were highly educated and paid much attention to news events. A majority of the participants had a university or an equivalent degree (56.9%). About half of the participants reported to read a newspaper every day (49.9%), about half watched the news on television every day (50.1%), and about a third of the participants both read a newspaper and watched the news on television every day (31.5%).

Materials

In the YNMT, the participants received 15 open-ended and 15 multiple-choice questions. These questions were selected semirandomly from a database consisting of 239 questions about news events that occurred from 1950 to 2006. The questions were taken from the AMV (Klomps, 2001) and the Daily News Memory Test (Meeter et al., 2005). Thirty-three news events came from the 1950s, forty from the 1960s and 1970s, thirty-seven from the 1980s, thirty-five from the 1990s, and fifty-four from the last 6 years. The events were categorised with regard to topic. The events involved international news ($n = 84$), entertainment news ($n = 44$), disasters ($n = 27$), Dutch national news ($n = 24$), sports news ($n = 21$), science news ($n = 17$), political news ($n = 15$), and crimes ($n = 7$). We have given the translations of a selection of 60 questions and their correct answers in Appendix A.

We programmed a dynamic selection procedure on trial level to obtain enough results for each of the three periods (i.e., younger than 10 years, between 10 and 25 years old, and older than 25 years), without making the test too difficult for younger participants or excluding older participants. Each question presented during the test by a certain participant had a 75% chance of being selected from the period after the participant's tenth birthday as calculated from the birth date given by the participant. If the question was not selected from the period after the participant's tenth birthday, then it was randomly selected from the entire period (i.e., 1950–2006).

For each question, three lures were created. The four possible answers were presented in random order. Participants had to select one alternative before they could proceed to next question (i.e., four-alternative forced-choice). No "I don't know" option was provided. Scoring answers to open-ended questions occurred automatically by matching the participant's answer against a word or a partial word indicative of the correct answer. Spelling mistakes or typing errors were neutralised by also matching on variants of the correct spelling. Eighteen questions were presented in multiple-choice format only; the others were presented as open-ended or multiple-choice questions.

Procedure

Before participants can participate in experiments on our website, they have to register. We ask for their gender, date of birth, level of education, and how many times a week they read a newspaper and watch the news. Furthermore, they have to choose a username and password that they can use to log in on other tests on our website without reregistering.

After registering or logging in, participants read a statement emphasising that the study was genuine and serious. Participants were asked to provide their informed consent, and were given instructions for the open-ended questions. Then, the participants were presented with 15 open-ended questions one by one. Participants had to enter the answer to the question in an open text field. When participants did not know the answer to the question, they were advised to take a guess. When they really did not know the answer, they had to enter a question mark before they could proceed to the next question. After the participants answered the open-ended questions, they were given the instructions for the multiple-choice questions. The 15 multiple-choice questions were also presented one by one. Participants were required to select one alternative before they could proceed. During the test, participants saw a progression bar at the top of the screen indicating how many questions they had already answered.

When the participants had answered the multiple-choice questions, they were thanked for their participation and informed about the purpose of this study. They were given a short summary of their results. They were told how many questions they had answered correctly, and how these results compared to the results of participants with similar background variables (i.e., gender, age, level of education, and news interest). Participants were also given the opportunity to ask questions or to give comments about the test.

The study was presented on the Internet. Besides numerous advantages to psychological research on the Internet, there are several serious problems (Reips, 2000, 2002). The first problem of Internet-based research is multiple submissions. However, one can minimise multiple submissions by supplying passwords. Participants had to register for the YNMT, but they were allowed to take the test as often as they wanted. We will only analyse the results of their first test. The results of subsequent tests will be dropped from the results. The second problem is self-selection. The self-selection bias can be controlled by the multiple site entry technique. We have promoted our website through other websites, search engines, traditional media, and word of mouth. The third possible problem of Internet-based research is the absence of a physical experimenter. The absence during the experiment could lead to problems when the instructions are unclear. Pretesting the experiment in a usability test and providing the possibility for feedback helps to improve the clarity of the instructions. The fourth problem concerning web-based research is the variance between computers, browsers, and networks, which could lead to reliability problems. However, extensive pretesting and random distribution of participants to experimental conditions in between-subjects designs reduces this problem. The use of within-subjects designs may eliminate this problem entirely. The final problem is the dropout rate. By giving financial incentives to participants who complete the experiment,

by giving immediate feedback about performance, by giving information about the duration of the experiment, and by giving feedback about the progress of the participant, dropout can be reduced.

Performance on the YNMT could be influenced by a second test on our website, the Galton-Crovitz test (Janssen et al., 2005, 2006; Janssen & Murre, 2007). Although it primarily deals with autobiographical memory, participants in the latter test are also asked to date 10 public events selected randomly from a database of 110 events. The 30 questions in the YNMT are the selected from a database of 239 questions, of which 80 covered events also present in the Galton-Crovitz database. Therefore, participants that completed both tests on average had 2.4 events (7.3% of 30) that featured in both tests (10 * 80/110 * 80/239). We looked at how many participants took both the Galton-Crovitz test and the YNMT within 24 hours. This was the case for 123 out of 1334, which means that just 0.67% of the events in the YNMT can be expected to also have been presented in the Galton-Crovitz test.

RESULTS

We omitted the results from tests that took less than 5 min ($N = 15$) or more than 30 min ($N = 45$) to complete, because it is possible that these tests were not taken seriously or that the correct answers to the questions were looked up. Without these omitted tests, participants took on average 11 min 25 s to complete the YNMT. We also excluded the results of questions that had average scores below .1 or above .9 (6 and 24 questions respectively) from the analyses, leaving a total of 209 questions. We will analyse the results of the multiple-choice and open-ended questions separately. We will first look at the results of the multiple-choice question (i.e., recognition), we will then look at the results of the open-ended questions (i.e., cued recall), and, finally, we will compare the results of the multiple-choice and open-ended questions.

Recognition

The average proportions correct and deviation scores of the multiple-choice questions as a function of the age at the event are displayed in Figure 1 per decade of questions. The average proportion correct of the multiple-choice questions was .720. The light-grey dashed lines represent questions about news events from the 1950s, while questions about recent news events (2000s) are represented by the black solid lines. The leftmost point of each line represents youngest participants in our study, whereas the rightmost point represents the oldest participants. In Panel A where the proportions correct are depicted, the recency effect is still visible. The proportion correct of
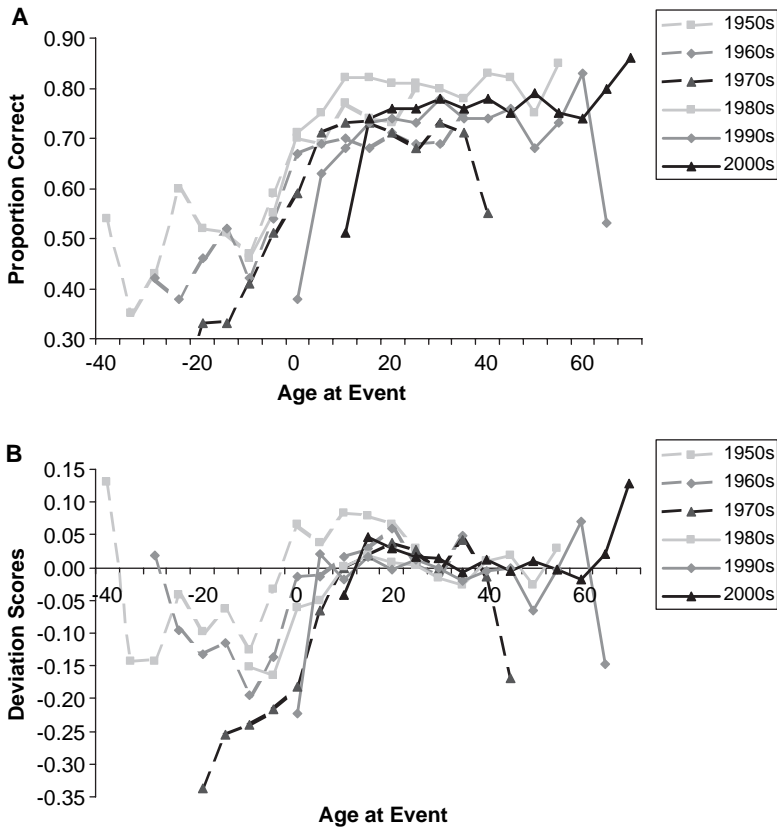
Figure 1. Average proportions correct (Panel A) and deviation scores (Panel B) of the multiple-choice questions for each decade of the questions as a function of the age at event. Age at event denotes the beginning of a 5-year range (e.g., "20" denotes age range 20–25).

questions about recent events are higher than the proportion correct of questions about events from other time periods, except for the proportion correct of questions about events from the 1980s.

We calculated a deviation score for each trial, assuming independent contributions of question and participant characteristics. The mean score of each question and each participant were used to compute an expectancy probability on each trial, which was subtracted from the observed score on the trial (0 or 1). The resulting deviation scores were then used in the statistical analyses below. The equations of the conditional probabilities approach are given in Appendix B with four examples.

We found no effect of gender, education, news interest, or age group on the deviation scores, nor did we find interactions between these factors.

These effects were not expected, since the mean score of each participant was factored out of the deviation score. Characteristics of the questions, such as topic and decade (e.g., recency effect), had no effect on the deviation scores for the same reason. However, we did find an interaction effect between age group and decade on deviation scores, $F(35, 17789) = 7.154$, $MSE = 0.138$, $p < .001$. Older adults perform better on questions about remote events than young adults, but this difference decreases as events become more recent.

The deviation scores of the multiple-choice questions are depicted in Panel B of Figure 1. For the statistical analysis, we divided the data into three lifetime periods. The events could have occurred before the participants' tenth birthday (before-10 events), between their tenth and twenty-fifth birthday (10–25 events), or after participants' twenty-fifth birthday (after-25 events). We found an overall effect of lifetime period, $F(2, 17822) = 64.302$, $MSE = 0.139$, $p < .001$. Questions about before-10 events ($M = -0.094$) were less often answered correctly than either those about 10–25 events ($M = 0.016$), $t(9392) = -10.837$, $p < .001$, or about after-25 events ($M = 0.002$), $t(10237) = -9.982$, $p < .001$. This result is not surprising, given that many of the before-10 events occurred before the participant was born. More interestingly, questions about 10-25 events were more often answered correctly than questions about after-25 events, $t(16015) = 2.448$, $p = .014$. Participants performed better on multiple-choice questions about public events that occurred in the period in which they were between 10 and 25 years old than on multiple-choice questions about public events that occurred in the period in which they were older than 25 years.

Cued recall

We have displayed the average proportions correct and deviation scores per decade as a function of the age at the event for the open-ended questions in Panels A and B of Figure 2. The average proportion correct of the open-ended questions was .465. We found no effect of gender, education, news interest, age group, topic, or decade on the deviation scores. Again, we did find an interaction effect between age group and decade, $F(35, 18669) = 7.804$, $MSE = 0.157$, $p < .001$. Older adults perform better on questions about remote events than young adults, but this difference decreases as events become more recent. We divided the data into the same three lifetime periods. There was an overall effect of lifetime period, $F(2, 18702) = 61.738$, $MSE = 0.157$, $p < .001$. Questions about before-10 events ($M = -0.083$) were answered correctly less often than those about 10–25 events ($M = 0.030$), $t(9716) = -10.893$, $p < .001$, or than those about above-25 events ($M = -0.006$), $t(10742) = -7.363$, $p < .001$. Questions about 10–25 events were answered more frequently correct than above-25 questions,
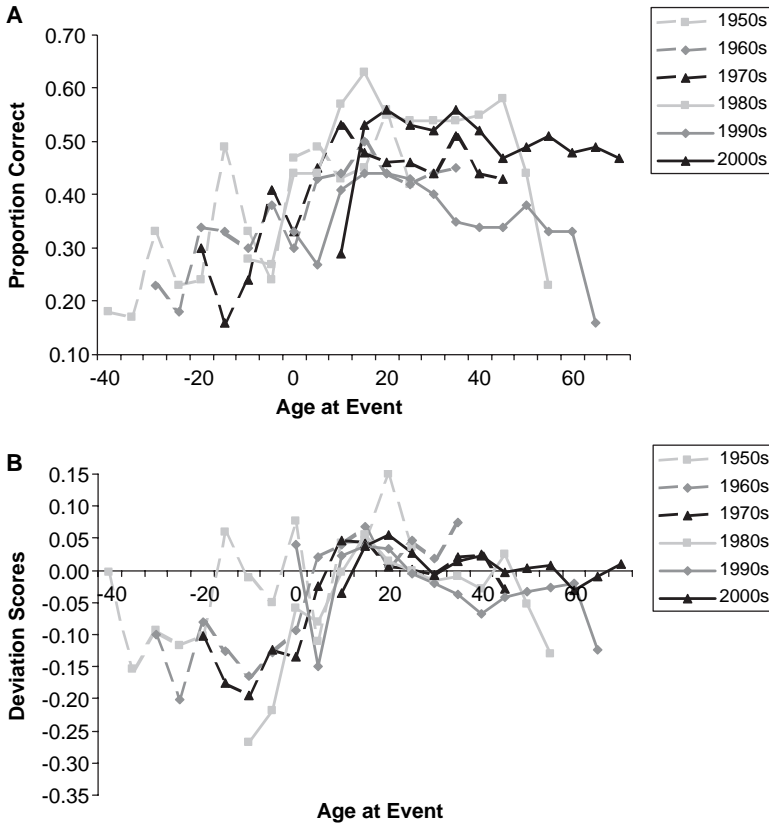
Figure 2. Average proportions correct (Panel A) and deviation scores (Panel B) of the open-ended questions for each decade of the questions as a function of the age at event. Age at event denotes the beginning of a 5-year range (e.g., "20" denotes age range 20–25).

$t(16946) = 5.888$, $p < .001$. Participants performed better on open-ended questions about public events that occurred in the period in which they were between 10 and 25 years old than on open-ended questions about public events that occurred in the period in which they were older than 25 years. The results thus mirrored the results of the multiple-choice questions.

Recognition versus cued recall

Finally, we looked whether the effect of lifetime period was stronger for open-ended questions than for multiple-choice questions. We did not find a main effect of question type ($p = .181$), but we did find an interaction effect between question type and lifetime period on the deviation scores,
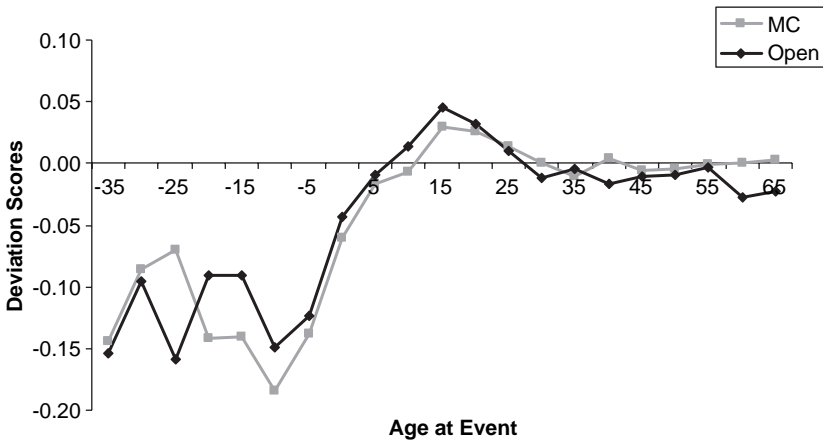
Figure 3. Average deviation scores for the multiple-choice (MC) and open-ended questions (Open) as a function of the age at event. Age at event denotes the beginning of a 5-year range (e.g., "20" denotes age range 20–25).

$F(5, 36369) = 3.398$, $MSE = 0.148$, $p = .033$. Participants performed relatively better on open-ended than on multiple-choice questions about events that had occurred before their twenty-fifth birthday and relatively worse on open-ended than on multiple-choice questions about events that had occurred after their twenty-fifth birthday. Figure 3 presents the average deviation scores of the multiple-choice and open-ended questions as a function of the age at events. We found that the open-ended questions were answered more frequently correct than the multiple-choice questions when the events came from the period in which the participants were between 10 and 25 years old, $t(15554) = -2.223$, $p = .026$. We found no difference between open-ended and multiple-choice questions when the events came from the period in which the participants were younger than 10 years ($p = .305$) or when the events came from the period in which the participants were older than 25 years ($p = .209$).

DISCUSSION

Here, we investigated the temporal distribution of memory for public events. Presenting the Yearly News Memory Test on the Internet gave us the opportunity to test our hypothesis in a large and diverse sample, thus increasing the external validity of the results (Reips, 2000, 2002). A caveat of using public items is that many news events have consequences that play out over time, making them less time-specific. Moreover, news events may be recalled in broadcasts and newspapers when similar events occur. The news

events can also be rehearsed through books or documentaries. We assumed that the aftermath for most participants had mainly strengthening effects, because the rehearsal of events has less impact and news coverage than the original event. Additional learning then reinforces memories that were already stored. However, from our data it is also clear that some people first learned about certain events years after they occurred. Some participants were able to answer a question correctly about an event that had occurred before their birth. Those people must have encoded new memories about these events. Such new memories, and indeed the strengthening of old ones, would act as a distortion of the results. This could have led to a flattening of the reminiscence bump in the temporal distribution of memory for public events.

We circumvented the issues of unequal item difficulty, intergenerational interest and matching by calculating deviation scores. This conditional probabilities approach takes into account independent effects of questions and participants on the likelihood that a question is answered correctly. We preferred this method to the standardisation of questions, because it does not rely on equivalent interest in categories of events across the decades and does not require matching of the groups on all possible participant characteristics. The method can adjust the scores of difficult or easy questions, such as questions about events that were less time-specific or questions about events from the 1980s, or the scores of participants with high or low scores. The recency effect, which could obscure the reminiscence bump in the lifetime distribution of young adults (Janssen et al., 2005), is also removed from the distributions with this method.

The results of the study show that the reminiscence bump can be found in the temporal distribution of memory for public events. Participants answered questions about events that occurred when they were between 10 and 25 years old more often correctly than questions about events that occurred before they were 10 years old or after they were 25 years old. Although the deviation scores of public events that occurred in the period in which participants were between 10 and 25 years old were the highest for events that had occurred in the 1950s (i.e., light-grey dashed lines in Figures 1 and 2), the results were not completely driven by those events from the 1950s. The trials about public events from 1950s only reflected a small portion of entire data set (4.8%), because about three-quarters of the questions were about events that had occurred after the participants' tenth birthday. The results of the study replicate those of Howes and Katz (1988), Rubin et al. (1998), and Schuman et al. (1997). However, we were able to circumvent issues, such as intergenerational interest and intergenerational matching, thus confirming the reminiscence bump in the temporal distribution of memory for public events.

Furthermore, we found that the reminiscence bump was larger in the results of the open-ended questions than in the results of the multiple-choice questions. Age at events had a larger effect on cued recall than on recognition. This difference cannot be explained by the fact that multiple-choice questions have higher proportions correct than open-ended questions, because the proportions correct are used separately to calculate the expected probabilities. This result is in line with episodic memory research that finds that age decrements are in general relatively smaller in tests of recognition than in tests of recall (e.g., Craik & McDowd, 1987; Nyberg et al., 2003; Spaan et al., 2003).

The reminiscence bump was thus found in a nonautobiographical memory domain. Memory for public events is not the first domain outside autobiographical memory in which the reminiscence bump has been found. The reminiscence bump also has been found in the distribution of favourite books, movies, records, and music genres (Holbrook & Schindler, 1989, 1996; Janssen et al., in press; Larsen, 1996; Rubin et al., 1998; Sehulster, 1996; Smith, 1994). A reminiscence bump in a nonautobiographical memory domain is predicted by the biological account that events are stored betted in adolescence and early adulthood, because the brain works at an optimum in those lifetime periods. The other accounts do not generate specific predictions about memory for public events, because they only refer to personal events: first-time experiences (cognitive account), self-defining moments (identity formation account), and transitional events (life scripts account). Although our results thus do not disprove any of these accounts, they constitute positive support only for the biological account.

One might argue, from the perspective of the identity formation and life scripts accounts, that people identify themselves more with public events from their teens than with public events from other lifetime periods, because adolescence and early adulthood provide a unique openness towards larger events (Schuman et al., 1997). Schuman and Scott (1989) found that people were more likely to mention an event as an event that changed the world when that event had happened when the participants were between 10 and 30 years old. However, most of the 200 public events in this study are unlikely to be considered world changing. Furthermore, it remains unclear whether people recall public events from their adolescence and early adulthood better, because they identify themselves more with those events, or whether people identify themselves with public events from their adolescence and early adulthood, because they have encoded those events better.

Again, though we take our results to support the biological account, they do not falsify other mechanisms, such as the cognitive account, because the four mechanisms are not mutually exclusive (Rubin et al., 1986, 1998), and each will have to be examined separately. However, the results of this and our previous studies; (Janssen et al., 2005; Janssen & Murre, 2007; Janssen

et al., in press) suggest that events may be encoded more strongly during adolescence and early adulthood, because the memory system is working more efficiently during that period, storing many public and personal events, including mundane personal events that are hardly stored earlier or later in life.

Although differential encoding may be the root cause of the reminiscence bump, we think that resampling may also have a role. Events from adolescence and early adulthood are resampled more frequently, because they have a larger likelihood to be recalled, they are self-defining memories, or they are transitional events. This makes them even stronger as time passes. Such resampling explains the increase in the reminiscence bump in the distribution of autobiographical memory of older adults (Janssen et al., 2005), the relative increase of remarkable events from adolescence as people become older (Janssen & Murre, 2007), and the larger reminiscence bump in the distribution of favourite records than the reminiscence bump in the distribution of favourite books (Janssen et al., in press).

The results do not explain why the brain works at an optimum in adolescence and early adulthood. Is this effect caused by changing levels of hormones or neurotransmitters? Or does working memory have a larger capacity in adolescence, enabling more events to be stored? More work, by psychologists as well as neuroscientists, will be required to answer this question.

REFERENCES

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, *2*(6), 396–408.

Berntsen, D., & Rubin, D. C. (2002). Emotionally charged autobiographical memories across the life span: The recall of happy, sad, traumatic, and involuntary memories. *Psychology and Aging*, *17*(4), 636–652.

Berntsen, D., & Rubin, D. C. (2004). Cultural life scripts structure recall from autobiographical memory. *Memory and Cognition*, *32*(3), 427–442.

Brewer, W. F. (1986). What is autobiographical memory? In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 25–49). Cambridge, UK: Cambridge University Press.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, *5*(1), 73–99.

Cermak, L. S. (1984). The episodic-semantic distinction in amnesia. In L. R. Squire & N. Butters (Eds.), *The neuropsychology of memory* (pp. 55–62). New York: Guilford Press.

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language*, *53*(4), 594–628.

Conway, M. A., & Haque, S. (1999). Overshadowing the reminiscence bump: Memories of a struggle for independence. *Journal of Adult Development*, *6*(1), 35–44.

Conway, M. A., Singer, J. A., & Tagini, A. (2004). The self and autobiographical memory: Correspondence and coherence. *Social Cognition*, *22*(5), 491–529.

Craik, F. I. M., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(3), 474–479.

Fitzgerald, J. M. (1988). Vivid memories and the reminiscence phenomenon: The role of a self narrative. *Human Development*, *31*, 261–273.

Holbrook, M. B., & Schindler, R. M. (1989). Some exploratory findings on the development of musical tastes. *Journal of Consumer Research*, *16*, 119–124.

Holbrook, M. B., & Schindler, R. B. (1996). Market segmentation based on age and attitude toward the past: Concepts, methods, and findings concerning nostalgic influences on customer tastes. *Journal of Business Research*, *37*, 27–39.

Holmes, A., & Conway, M. A. (1999). Generation identity and the reminiscence bump: Memory for public and private events. *Journal of Adult Development*, *6*(1), 21–34.

Howes, J. L., & Katz, A. N. (1988). Assessing remote memory with an improved public events questionnaire. *Psychology and Aging*, *3*(2), 142–150.

Hyland, D. T., & Ackerman, A. M. (1988). Reminiscence and autobiographical memory in the study of the personal past. *Journal of Gerontology*, *43*(2), 35–39.

Janssen, S. M. J., Chessa, A. G., & Murre, J. M. J. (2005). The reminiscence bump in autobiographical memory: Effects of age, gender, education, and culture. *Memory*, *13*(6), 658–668.

Janssen, S. M. J., Chessa, A. G., & Murre, J. M. J. (2006). Memory for time: How people date events. *Memory and Cognition*, *34*(1), 137–148.

Janssen, S. M. J., Chessa, A. G., & Murre, J. M. J. (in press). Temporal distribution of favourite books, movies and records: Differential encoding and re-sampling? *Memory*.

Janssen, S. M. J., & Murre, J. M. J. (2007). Reminiscence bump in autobiographical memory: Unexplained by first-time experiences or valence and emotionality of events. *Quarterly Journal of Experimental Psychology. Manuscript submitted for publication*.

Klomps, P. (2001). *Retrograde amnesie: Constructie van een vragenlijst* [*Retrogradeamnesia: Construction of a questionnaire*]. Unpublished research report, University of Amsterdam, The Netherlands.

Larsen, S. F. (1996). Memorable books: Recall of reading and its personal context. In R. Kreuz & M. S. MacNealy (Eds.), *Empirical approaches to literature and aesthetics* (pp. 585–599). Norwood, NJ: Ablex.

Longmore, F. J., Knight, R. G., & Longmore, B. E. (1990). A test of remote memory for use with New Zealand subjects. *New Zealand Journal of Psychology*, *19*(1), 17–23.

Meeter, M., & Murre, J. M. J. (2004). Consolidation of long-term memory: Evidence and alternatives. *Psychological Bulletin*, *130*(6), 843–857.

Meeter, M., Murre, J. M. J., & Janssen, S. M. J. (2005). Remembering the news: Modeling retention data from a study with 14,000 participants. *Memory and Cognition*, *3*(5), 793–810.

Merriam, S. B., & Cross, L. H. (1982). Adulthood and reminiscence: A descriptive study. *Educational Gerontology*, *8*(3), 275–290.

Murre, J. M. J. (1996). TraceLink: A model of amnesia and consolidation of memory. *Hippocampus*, *6*(6), 675–686.

Murre, J. M. J. (1997). Implicit and explicit memory in amnesia: Some explanations and predictions by the TraceLink model. *Memory*, *5*(1–2), 213–232.

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

Niedźwieńska, A. (2003). Gender differences in vivid memories. *Sex Roles*, *49*(7–8), 321–331.

Nyberg, L., Maitland, S. B., Rönnlund, M., Bäckmann, L., Dixon, R. A., Wahlin, Å., et al. (2003). Selective adults age differences in an age-invariant multifactor model of declarative memory. *Psychology and Aging*, *18*(1), 149–160.

Pillemer, D. B. (2001). Momentous events and the life story. *Review of General Psychology*, *5*(2), 123–134.

Reips, U.-D. (2000). The web experimenting method: Advantages, disadvantages, and solutions. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 89–117). San Diego, CA: Academic Press.

Reips, U.-D. (2002). Standards for Internet-based experimenting. *Experimental Psychology*, *49*(4), 243–256.

Robinson, J. A. (1976). Sampling autobiographical memory. *Cognitive Psychology*, *8*(4), 578–595.

Robinson, J. A. (1986). Autobiographical memory: A historical prologue. In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 19–24). Cambridge, England: Cambridge University Press.

Robinson, J. A. (1992). First experience memories: Context and functions in personal histories. In M. A. Conway, D. C. Rubin, H. Spinnler, & W. A. Wagenaar (Eds.), *Theoretical perspectives on autobiographical memory* (pp. 223–236). Dordrecht, The Netherlands: Kluwer Academic.

Rosenbaum, R. S., Winocur, G., & Moscovitch, M. (2001). New views on old memories: Re-evaluating the role of the hippocampal complex. *Behavioral Brain Research*, *127*(1–2), 183–197.

Rubin, D. C., & Berntsen, D. (2003). Life scripts help to maintain autobiographical memories of highly positive, but not highly negative, events. *Memory and Cognition*, *31*(1), 1–14.

Rubin, D. C., Rahhal, T. A., & Poon, L. W. (1998). Things learned in early adulthood are remembered best. *Memory and Cognition*, *26*(1), 3–19.

Rubin, D. C., Schulkind, M. D., & Rahhal, T. A. (1999). A study of gender differences in autobiographical memory: Broken down by age and sex. *Journal of Adult Development*, *6*(1), 61–71.

Rubin, D. C., Wetzler, S. E., & Nebes, R. D. (1986). Autobiographical memory across the adult life span. In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 202–221). Cambridge, UK: Cambridge University Press.

Rybash, J. M. (1999). Aging and autobiographical memory: The long and bumpy road. *Journal of Adult Development*, *6*(1), 1–10.

Rybash, J. M., & Monaghan, B. E. (1999). Episodic and semantic contributions to older adults' autobiographical recall. *Journal of General Psychology*, *126*(1), 85–96.

Schooler, L. J., Shiffrin, R. M., & Raaijmakers, J. G. W. (2001). A Bayesian model for implicit effects in perceptual identification. *Psychological Review*, *108*(1), 257–272.

Schrauf, R. W., & Rubin, D. C. (2001). Effects of voluntary immigration on the distribution of autobiographical memory over the lifespan. *Applied Cognitive Psychology*, *15*(7), S75–S88.

Schuman, H., Belli, R. F., & Bischoping, K. (1997). The generational basis of historical knowledge. In J. W. Pennebaker, D. Paez, & B. Rimé (Eds.), *Collective memory of political events* (pp. 47–77). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Schuman, H., & Scott, J. (1989). Generations and collective memories. *American Sociology Review*, *54*(3), 359–381.

Sehulster, J. R. (1996). In my era: Evidence for the perception of a special period of the past. *Memory*, *4*(2), 145–158.

Smith, T. W. (1994). Generational differences in musical preferences. *Popular Music and Society*, *18*, 43–59.

Spaan, P. E. J., Raaijmakers, J. G. W., & Jonker, C. (2003). Alzheimer's disease versus normal ageing: A review of the efficiency of clinical and experimental memory measures. *Journal of Clinical and Experimental Neuropsychology*, *25*(2), 216–233.

Squire, L. R. (1974). Remote memory as affected by aging. *Neuropsychologica*, *12*(4), 429–435.

Squire, L. R., & Slater, P. C. (1975). Forgetting in very long-term memory as assessed by an improved questionnaire technique. *Journal of Experimental Psychology: Human Learning and Memory*, *1*(1), 50–54.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York: Academic Press.

Warrington, E. K., & Sanders, H. I. (1971). The fate of old memories. *Quarterly Journal of Experimental Psychology*, *23*(4), 432–442.

Webster, J. D., & McCall, M. E. (1999). Reminiscence functions across adulthood: A replication and extension. *Journal of Adult Development*, *6*(1), 73–85.

APPENDIX A: TRANSLATIONS OF A SELECTION OF QUESTIONS AND THEIR CORRECT ANSWERS

Question	Answer
What was the name of the American senator, who accused many people of Communist sympathies?	Joseph McCarthy
Why did the first administration of Prime Minister Drees resign in 1951?	Over Papua New Guinea
Who succeeded Stalin as leader of the Soviet Union in 1953?	Nikita Khrushchev
In which African country was the Mau Mau rebel movement active that attacked white settlers in 1953?	Kenya
Who were the first climbers to reach the summit of Mount Everett in 1953?	Edmund Hillary & Tenzing Norgay
What was the name of the Afro-American woman, who refused to give up her seat in the bus in 1955?	Rosa Parks
Who developed in 1955 the first polio vaccine?	Jonas Salk
Who was the first female minister in The Netherlands?	Marga Klompé
What was the name of the dog, which became the first mammal in space in 1957?	Laika
Who was president of Cuba, before the revolution led by Fidel Castro in 1959?	Fulgencio Batista
What was the name of the first Prime Minister of Congo, who was killed in 1961?	Patrice Emery Lumumba
What was the name of the Secretary-General of the United Nations, who died in a plane crash in 1961?	Dag Hammarskjöld
In which city was US President John F. Kennedy assassinated in 1963?	Dallas
What was the original name of the boxer Muhammad Ali, who became world champion for the first time in 1964?	Cassius Clay
In which country did the Tet offensive take place?	Vietnam
As what kind of animal did author Gerard Reve present God, for which he was prosecuted for blasphemy?	Donkey
Who killed Dr Martin Luther King, Jr in 1968?	James Earl Ray

What was the name of the Czechoslovakian student, who set himself on fire in 1969?	Jan Palach
Who were the first two astronauts on the moon in 1969?	Neil Armstrong & Buzz Aldrin
Who crossed the Atlantic Ocean in a papyrus boat in 1970?	Thor Heyerdahl
Which country was declared independent from Pakistan in 1971?	Bangladesh
How are the riots in Londonderry in 1972 also called?	Bloody Sunday
What was the name of the bloodless military coup in Portugal in 1974?	Carnation Revolution
Who was the leader of the People's Temple, a religious cult best known for its mass suicide in Guyana in 1978?	Jim Jones
What was the name of the Ugandan dictator, who was exiled in 1979?	Idi Amin Dada
Who was the last Viceroy of India, who was killed by the IRA in 1979?	Lord Mountbatten
Which Swedish tennis player won Wimbledon for the fifth time in a row in 1980?	Björn Borg
Which country attacked the British Falkland Islands in 1981?	Argentina
Which country won the World Cup football in 1982?	Italy
In which country did the December killings take place in 1982?	Suriname
In which Indian city did poisonous gas escape from a pesticide plant in 1984, killing more than 20,000 people?	Bhopal
Which Space Shuttle exploded briefly after its launch in 1986?	Challenger
What was the name of the ferryboat capsized in the English Channel in 1987?	*Herald of Free Enterprise*
What was the name of the jazz musician who died in 1988 after he fell out of a window of a hotel in Amsterdam?	Chet Baker
In which Scottish town did an American airplane crash in 1988?	Lockerbie
What was the name of the tanker, which spilled oil on the coast of Alaska in 1989?	*Exxon Valdez*
Who became the first Chancellor of a unified Germany in 1990?	Helmut Kohl
What was the name of the Afro-American man, who was beaten up by four police officers in Los Angeles in 1991?	Rodney King
From which disease did Freddie Mercury, who was the singer of the rock group Queen, die in 1991?	AIDS
Which rock singer committed suicide in 1994?	Kurt Cobain
Which Japanese city was damaged heavily by an earthquake in 1995?	Kobe

In which city was Sarin gas used in an attack on the subway in 1995?	Tokyo
In which American city was a government building destroyed by a bomb attack in 1995?	Oklahoma City
What was the name of the chess computer, which defeated world-champion Garry Kasparov in 1996?	Deep Blue
Which fashion designer was killed in Florida in 1997?	Gianni Versace
How did Princess Diana die in 1997?	Car accident
What was the name of the intern, with whom president Clinton had a sexual relationship?	Monica Lewinsky
Which famous discotheque in Amsterdam burned down completely in 1999?	Roxy
In which city did a firework depot explode in 2000?	Enschede
What was the name of the Russian submarine, which sank in the Barents Sea in 2000?	*Kursk*
From which building did Herman Brood jump to commit suicide in 2001?	Hilton Hotel
Which rare, but deadly disease was found in mail in the United States in 2001?	Anthrax
In which city was Pim Fortuyn assassinated in 2002?	Hilversum
On which Indonesian island were 202 people killed in a bomb attack in 2002?	Bali
What was the name of the UN weapons inspector, who did not find any weapons of mass destruction in Iraq in 2002?	Hans Blix
Who won the first season of *Idols*?	Jamaï
Of which country was Anna Lindh, who was assassinated in 2003, Secretary of State?	Sweden
Which singer showed her nipple during the halftime entertainment of the Super Bowl in 2004?	Janet Jackson
On which island did the American girl Natalee Holloway disappear?	Aruba
What was the name of the hurricane that flooded New Orleans in 2005?	Katrina

APPENDIX B

To calculate a deviation score of a trial in which subject i answers question x, we computed for each trial the likelihood that subject i answers question x correctly. We derived this likelihood from the odds in favour of a correct answer. The difference between probabilities and odds can be illustrated with the rolling of a die. The probability of rolling the number three is 1/6, but the odds in favour of rolling the number three are 1/5, since one outcome is the number three and five outcomes are not. It is fairly easy to compute the odds from the probability and vice versa with the two following equations, which are also the basis of Equations 3, 4, 5, and 8 below:

(1) odds = probability/[1 − probability]
(2) probability = odds/[1 + odds]

We will first give the equations to calculate the deviation score and then give four worked examples. We compute first the odds in favour of a correct answer when a random subject answered a random question, $odds(c)$ (Equation 3). Then, we calculate for each subject the odds in favour of a correct answer when subject x answered a random question and for each question the odds in favour of a correct answer when a random subject answered question i, $odds(c \mid x)$, and $odds(c \mid i)$ (Equations 4 and 5). In the equations below, $p(c)$ stands for the likelihood that a random question is answered correctly by a random subject, $p(c \mid x)$ for the likelihood that subject x answers a random question correctly, and $p(c \mid i)$ for the likelihood that a random subject answers question i correctly.

(3) odds(c) = p(c)/[1 − p(c)]
(4) odds(c | x) = p(c | x)/[1 − p(c | x)]
(5) odds(c | i) = p(c | i)/[1 − p(c | i)]

The odds per subject and per question can be combined into the odds in favour of correct answer when subject x answers question i, $odds(c \mid x \, \& \, i)$, under the assumption that subject characteristics and item difficulty contribute independently to the odds in favour of answering a question correctly. This assumption can be interpreted as:

(6) odds(c | x & i)/odds(c | i) = odds(c | x)/odds(c)

That is, the change in odds from when it is given that question i is answered correctly to that is given that subject x answers question i correctly should be the same as the change in odds from when a random person answers a random question correctly to when it is given that subject x

answers a random question correctly. Equation 7 for *odds(c | x & i)* can be derived from Equation 6. Then, we calculated for each trial the expected probability using the odds (Equation 8). Finally, these expected probabilities were subtracted from the observed score, $s_{i,x}$ (equal to 0 or 1), of the trials to yield deviation scores, $d_{i,x}$ (Equation 9).

(7) odds(c | x & i) = [odds(c | x) * odds(c | i)]/odds(c)
(8) p(c | x & i) = odds(c | x & i) / [1 + odds(c | x & i)]
(9) $d_{i,x}$ = $s_{i,x}$ − p(c | x & i)

Suppose that on 60% of all trials a correct answer is given. Subject *x* has a high average score, answering 80% of the questions correctly. Subject *y* has a low average score, answering 40% correctly. Difficult question *i* is answered correctly by 40% of the subjects and easy question *j* is answered correctly by 80%. Below we have calculated the odds, using the formulas above, of all trials (Equation 3), the subjects *x* and *y* (Equation 4), and the questions *i* and *j* (Equation 5).

(3) odds(c) = 0.6/[1 − 0.6] = 0.6/0.4 = 1.5
(4) odds(c | x) = 0.8/[1 − 0.8] = 0.8/0.2 = 4,
    odds(c | y) = 0.4/[1 − 0.4] = 0.4/0.6 = 0.667
(5) odds(c | i) = 0.4/[1 − 0.4] = 0.4/0.6 = 0.667,
    odds(c | j) = 0.8/[1 − 0.8] = 0.8/0.2 = 4

We have then calculated the odds for trials when subject *x* answers question *i*, subject *x* answers question *j*, subject *y* answers question *I*, and subject *y* answers question *j* correctly (7). Subsequently, we have calculated the expected probabilities of a correct answer using Equation 8. A trial in which question *i* which was answered correctly by subject *y* would have had higher deviation score than a trial in which question *j* was answered correctly by subject *y*, or a trial in which question *i* or question *j* was answered correctly by subject *x* (9). The opposite is true for trials that were answered incorrectly. Thus, trials of subjects with a low average score answering difficult questions correctly or trials of subjects with a high average score answering easy questions incorrectly have a bigger impact on the results than other trials.

(7) odds(c | x & i) = [4 * 0.667]/1.5 = 2.667/1.5 = 1.778,
    odds(c | x & j) = [4 * 4]/1.5 = 16/1.5 = 10.667,
    odds(c | y & i) = [0.667 * 0.667]/1.5 = 0.444/1.5 = 0.297,
    odds(c | y & j) = [0.667 * 4]/1.5 = 2.667/1.5 = 1.778
(8) p(c | x & i) = 1.778/[1 + 1.778] = 1.778/2.778 = 0.640,
    p(c | x & j) = 10.667/[1 + 10.667] = 10.667/11.667 = 0.914,

$p(c \mid y \ \& \ i) = 0.297/[1 + 0.297] = 0.297/1.297 = 0.229,$

$p(c \mid y \ \& \ j) = 1.778/[1 + 1.778] = 1.778/2.778 = 0.640$

(9) For correctly answered trials:

$d_{x,i} = 1 - 0.640 = 0.360,$

$d_{x,j} = 1 - 0.914 = 0.086,$

$d_{y,i} = 1 - 0.229 = 0.771,$

$d_{y,j} = 1 - 0.640 = 0.360$

(9) For incorrectly answered trials:

$d_{x,i} = 0 - 0.640 = -0.640,$

$d_{x,j} = 0 - 0.914 = -0.914,$

$d_{y,i} = 0 - 0.229 = -0.229,$

$d_{y,j} = 0 - 0.640 = -0.640$