

# Thinking about memories for everyday and shocking events: Do people use ease-of-retrieval cues in memory judgments?

GERALD ECHTERHOFF

*University of Cologne, Cologne, Germany*

and

WILLIAM HIRST

*New School University, New York, New York*

Extant research shows that people use retrieval ease, a feeling-based cue, to judge how well they remember life periods. Extending this approach, we investigated the role of retrieval ease in memory judgments for single events. In Experiment 1, participants who were asked to recall many memories of an everyday event (New Year's Eve) rated retrieval as more difficult and judged their memory as worse than did participants asked to recall only a few memories. In Experiment 2, this ease-of-retrieval effect was found to interact with the shocking character of the remembered event: There was no effect when the event was highly shocking (i.e., learning about the attacks of September 11, 2001), whereas an effect was found when the event was experienced as less shocking (due either to increased distance to "9/11" or to the nonshocking nature of the event itself). Memory vividness accounted for additional variance in memory judgments, indicating an independent contribution of content-based cues in judgments of event memories.

Memory research has not only investigated the processes and contents of remembering, but also people's beliefs and judgments about how and what they remember (see, e.g., Herrmann, 1982; Koriat, 1993; Metcalfe, 2000; Nelson, 1992). A central memory judgment is the assessment of the quality of one's recollections. When people judge the quality of their memories, they estimate how well or how accurately they remember past experiences. When people believe that their memories are accurate and reliable, they will be more willing to draw conclusions from the remembered experiences and act accordingly (see Koriat & Goldsmith, 1996). In contrast, when people suspect that their recall is not accurate or reliable, they are more susceptible to social influence (e.g., Bless, Strack, & Walther,

2001). How do people arrive at such judgments? What information do they use to assess how well they remember a target event or episode? Two main classes of cues have been described: *content-based* and *feeling-based* cues (see Koriat & Levy-Sadot, 1999; Whittlesea & Leboe, 2000). On the one hand, people may consider the content of a retrieved memory, such as its vividness or visual details, to determine how well and how reliably they remember the target information (e.g., Conway, 1996; Johnson & Raye, 1981; Ross, Buehler, & Karr, 1998; Ross & MacDonald, 1997). On the other hand, people may base their memory judgments on feelings concerning their retrieval performance, such as the experienced ease of retrieving relevant information (e.g., Belli, Winkielman, Read, Schwarz, & Lynn, 1998; Benjamin, Bjork, & Schwartz, 1998; Kelley & Lindsay, 1993; Winkielman, Schwarz, & Belli, 1998). Whereas the use of retrieval ease has been demonstrated for entire life periods such as one's childhood (Belli et al., 1998; Winkielman et al., 1998), it is unclear whether rememberers refer to such a feeling-based cue in memory judgments for single events. The present research was conducted to investigate this open issue: Would rememberers use retrieval ease in judgments about memories for single events (i.e., episodic memories), even though such memories are characteristically different from memories for life periods?

There is little doubt that content-based cues can guide judgments about episodic or event memories. In content-based memory judgments, the rememberer takes into account attributes of the memory itself, such as its vividness

---

G.E. is now at the University of Bielefeld. This research was supported by NIMH Grant MH066972 to W.H. We thank Hartmut Blank, Gerd Bohner, Bertram Gawronski, and Walter Hussy for providing valuable comments on earlier drafts of this article. We also extend our gratitude to Marlene Karst, Lynn Kummerfeldt, Kim Rohwer, and Megan Rudesill for their dedicated assistance in data collection and administration, and to Robert Meksin and Birgit Neumann for their support in the preparation of the manuscript. Parts of this research were presented at the 5th Meeting of the Society for Applied Research in Memory and Cognition, held at Aberdeen University in July 2003. Correspondence concerning this article should be addressed to G. Echterhoff, Abteilung für Psychologie, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany (e-mail: gerald.echterhoff@uni-bielefeld.de).

*Note*—This article was accepted by the previous editorial team, when Colin M. MacLeod was Editor.

or perceptual details (Conway, 1996; Conway & Pleydell-Pearce, 2000; Johnson, Foley, Suengas, & Raye, 1988; Johnson, Hashtroudi, & Lindsay, 1993; Johnson & Raye, 1981; Ross & MacDonald, 1997). Vivid imagery seems to play a central role in such judgments (see Brewer, 1986; Conway, 1990; Conway & Pleydell-Pearce, 2000; Garry, Manning, Loftus, & Sherman, 1996; Rubin, 2005; Williams, Healy, & Ellis, 1999). For instance, Dewhurst and Conway (1994) demonstrated that memories are experienced as subjectively “real” when they contain vivid visual images.

However, content attributes are not the only information from which people may infer the quality of their memories. People know that some pieces of information come to mind more easily than others, and they may use such feelings of ease to make a variety of memory judgments (see, e.g., Benjamin & Bjork, 1996; Kelley & Lindsay, 1993; Koriat & Levy-Sadot, 1999; Whittlesea & Leboe, 2000). Thus, people may derive judgments not only from *what* comes to mind but also from *how* that information comes to mind. Early work on the tip-of-the-tongue experience (e.g., Hart, 1965) indicated that people may have a more or less pronounced feeling that they will be able to recall a specific piece of information. A burgeoning body of subsequent research, typically subsumed under the field of metacognition, has explored in more detail such feelings of knowing or intuitions about successful future retrieval (e.g., Metcalfe, 2000; Nelson, 1992). More specifically, researchers have identified the ease of processing or retrieving pieces of knowledge (see Schwarz et al., 1991; Tversky & Kahneman, 1973) as a potent cue for judgments of future memory performance (e.g., Begg, Duft, Lalonde, Melnick, & Sanvito, 1989; Benjamin & Bjork, 1996; Benjamin et al., 1998; Koriat, 1993; Mazzoni & Nelson, 1995). The diagnosticity of retrieval ease in making judgments about memory performance is supported by research revealing faster retrieval and greater feelings of knowing for more overlearned material (e.g., Nelson, Leonesio, Shimamura, Landwehr, & Narens, 1982).

Whereas these studies have revealed the use of feeling-based cues for assessments of one’s future memory performance, relatively little research has applied this rationale to memory judgments at the time of retrieval. Kelley and Lindsay (1993) found that prior exposure to answers to general knowledge questions increased not only the speed but also the confidence with which participants gave the primed answers on a knowledge test. According to a widely accepted interpretation, the priming procedure enhanced the ease of retrieving an answer from memory, which, in turn, induced people to feel more confident in the accuracy of their answers (for research revealing similar effects on other types of judgments, see, e.g., Arkes, Hackett, & Boehm, 1989; Hasher, Goldstein, & Toppino, 1977).

Subsequent studies by Winkielman et al. (1998) and Belli et al. (1998) extended the ease-of-retrieval approach to the domain of personal memory, demonstrating that people may use the experienced ease of retrieving specific events for judgments of memory related to an entire

period of their lives—that is, childhood after the age of 5 (see also Skurnik, Schwarz, & Winkielman, 2000). In these studies, participants rated their childhood memory as worse when they had retrieved many (e.g., 12 or 8) childhood events prior to their memory judgment than when they had recalled only a few (i.e., 4) events. Since participants reported experiencing the retrieval of many events as more difficult than the retrieval of a few, the researchers argued that participants based their memory judgments on the experienced ease of retrieving events from the target episode (for similar findings in other judgment domains, see, e.g., Haddock, Rothman, Reber, & Schwarz, 1999; Rothman & Schwarz, 1998; Schwarz et al., 1991; Wänke, Bohner, & Jurkowsch, 1997; for a review, see Schwarz & Vaughn, 2002).

None of this research has specifically examined memories for personally experienced single events. Event memories are more restricted in temporal extension and less general than memories for life periods (Anderson & Conway, 1993; Conway, 1996; Conway & Pleydell-Pearce, 2000), and they typically consist of more specific actions or subepisodes (see, e.g., Anderson & Conway, 1993; Reiser, Black, & Abelson, 1985; Reiser, Black, & Kalamarides, 1986). People may experience greater difficulty retrieving subcomponents of single events when required to recall many of them, and as a result, they may judge their event memory to be poor. On the other hand, event memories may differ categorically from memories for life periods. In event memory, the recall of specific microevents or details may inadvertently focus a rememberer’s attention on content-based cues such as visual details. These content-based cues may, in turn, outweigh any concomitant feeling-based cues. In contrast, memories for entire life periods such as one’s childhood merely require the retrieval of rather abstract or general events (e.g., attending school), for which the recall of content-based cues (such as sensory information) is not only unnecessary, but may not even be possible.<sup>1</sup> Experiment 1 was conducted to shed light on this issue, investigating whether people would use ease of retrieval in memory judgments for personally experienced single events.

Even if retrieval ease is found to affect memory judgments for such events, one may suspect that this effect does not necessarily apply to all kinds of event memories. Specifically, there is one class of event memories that seems to impose great constraints on memory judgments—that is, memories for publicly shocking events, such as murders of political leaders, natural disasters, or major terrorist attacks (see Brown & Kulik, 1977; Neisser & Harsch, 1992; Schmidt, 2004; Talarico & Rubin, 2003; Weaver, 1993). The critical difference between such “flashbulb memories” and memories for mundane, nonshocking events is not found in their actual accuracy or quality, but rather in the concomitant judgments of memory quality: Flashbulb memories are apparently no more accurate than other episodic memories (McCloskey, Wible, & Cohen, 1988; Talarico & Rubin, 2003; Wright, 1993), but subjectively they appear exceptionally accurate and faithful. Although

the exact mechanism underlying such memory judgments is not known (see Talarico & Rubin, 2003), there is reason to suspect that these judgments are due to factors other than ease. Flashbulb events are usually vividly remembered (see Rubin & Kozin, 1984). Moreover, people often hold a belief that memories such as those of the terrorist attacks on September 11, 2001, will never be forgotten (see, e.g., the Web site [www.wwnfspt11.com](http://www.wwnfspt11.com)). It is quite possible that these distinctive characteristics of flashbulb memories limit any influence ease of retrieval may have on memory quality judgments.

Support for this proposal can be found in the literature on judgments outside the memory domain. It indicates that high personal involvement in a nonmnemonic judgment task (Rothman & Schwarz, 1998) or extreme attitudes about the judgment object (Haddock et al., 1999) may preclude the use of ease-of-retrieval cues. For instance, Haddock et al. demonstrated that whereas participants with a moderate attitude concerning an issue (doctor-assisted suicide) used the subjective ease of retrieving issue-related arguments to infer their attitude, participants holding extreme attitudes did not. Haddock et al. argued that extreme-attitude participants did not refer to retrieval ease because they had already developed a chronically accessible evaluation of the attitude object that they used to form their judgments, irrespective of their retrieval feelings. Although these results were obtained for judgments of attitudes, they may still point to similar processes in the domain of memory judgments: When people are still highly concerned about a past event or when they experience high emotional involvement with an event, they may readily derive their memory judgments without reference to ease of retrieval.

Thus, after demonstrating the use of ease of retrieval for a socially synchronized, nonshocking event (Experiment 1), we will compare memory judgments for the same nonshocking event with those for a shocking public event, "9/11" (Experiment 2).

## EXPERIMENT 1

In addition to extending the findings of Belli et al. (1998) and Winkielman et al. (1998) to judgments about episodic or event memory, we also introduced some additional methodology and measures. First, the analyses reported by Belli et al. and Winkielman et al. did not explicitly address the mediating role of experienced ease of retrieval. These authors demonstrated both that the experimental manipulation (recalling few vs. many events) affected experienced ease and that ease was correlated with memory judgments (Belli et al., 1998). However, the correlation between ease and memory judgments did not control for the recall task manipulation. We employed a mediational analysis (see Baron & Kenny, 1986) to examine the relations among the independent variable (the quantity of recalled memories), the proposed mediator (retrieval ease), and the dependent variable (memory judgment), including the simultaneous effects of both memory quantity and retrieval ease in predicting memory judgments.

Second, it is conceivable that the ease-of-retrieval manipulation employed by Belli et al. (1998) and Winkielman et al. (1998) not only affected experienced ease but also content-based cues. Possibly, memories provided toward the end of the recall list in the many-events condition were not as vivid nor as rich in other attributes as memories provided in the few-events condition. In order to explore this possibility, we included ratings of a critical content-based cue (vividness), in addition to ratings of experienced ease.

In this experiment, participants were asked to rate the quality of their memories for their experiences on the previous New Year's Eve, after an interval of either 1 or 6 months. The experimental manipulation followed the procedure employed by Winkielman et al. (1998) and Belli et al. (1998), who instructed participants to recall either few or many instances from a target period prior to their judgments. The recall of fewer single memories should be experienced as relatively easy, whereas the recall of a larger number of memories should be experienced as rather difficult. If participants based their memory judgments on experienced ease of retrieval (a feeling-based cue), ratings of memory quality should be lower after the recall of many rather than few specific memories.

## Method

**Participants.** The participants were 93 people who responded to advertisements placed at the New School University in New York City (49 students of psychology, 23 students of other disciplines, 20 nonstudents; 57 female, 35 male; mean age = 25.2 years; 1 participant did not report demographic information). Forty-four of them participated 1 month after the target event (New Year's Eve), and the remaining 49 participated after an interval of 6 months. None of the demographic variables were significantly correlated with the dependent measures. The psychology students among the participants received curricular credit for their participation; all others received no compensation.

**Design.** The study was based on a 2 (number of memories: few vs. many)  $\times$  2 (delay: 1 month vs. 6 months) between-participants design. The participants were randomly assigned to the few- or many-memories condition. The main dependent variable consisted of their quality judgments concerning their memories for experiences on the previous New Year's Eve (December 31).

**Materials and Procedure.** All participants were randomly handed one of two versions of a test booklet titled "Memory Questionnaire." They were told to answer all questions in the given order. Those in the many-memories condition were first asked to recall 12 memories of their experiences on the previous New Year's Eve. In the few-memories condition, the participants were asked to recall 4 memories. All participants were told not to include general facts about New Year's Eve. They wrote down the different memories on separate lines below the task instruction, which were numbered according to the number of required memories. On the next page of the test booklet, the participants were asked to judge the quality of their memory for the previous New Year's Eve: "How good is your memory for what you did or experienced last New Year's Eve?" (1 = *very poor*, 9 = *very good*) On the same page, they rated the vividness of their memory (1 = *not vivid at all*, 9 = *very vivid*). On a separate page, the participants then rated the difficulty of the recall experience ("How difficult was the above memory task for you?"; 1 = *very easy*, 9 = *very difficult*). Those in the few-memories condition were also asked to recall another 8 memories on an additional page. Finally, on the last page of the test booklet, all participants indicated what they perceived to be the purpose and rationale of the study.

**Assessment of recall performance.** Although we carefully phrased the instructions for the recall task to elicit specific episodic memories from the participants, we also conducted a content analysis to determine whether the responses fit the criterion. We defined as acceptable those responses that referred to personal experiences during the target episode (the previous New Year's Eve). An acceptable response from the participants' response sheets was, for instance, "I talked to my mom on the mobile phone shortly after midnight." The coding manual also defined cases of unacceptable responses, such as general facts (i.e., impersonal knowledge) about New Year's Eve (e.g., "Usually, there are fireworks") or memories falling outside the required time frame (e.g., "I prepared food for the party 2 days before"). Two independent coders first eliminated redundant responses for each participant and then categorized the remaining responses as either acceptable or unacceptable (with sufficiently high intercoder agreement,  $\kappa = .81$ ). Discrepancies between the two coders were resolved with the help of a third coder. Three participants were excluded from the analyses because their recall reports contained a large number of unacceptable responses (from six to eight). These 3 people are not included in the 93 participants described in the Participants section above.

## Results and Discussion

**Manipulation checks.** A manipulation check for the number of recalled memories confirmed that participants in the many-memories condition recalled significantly more memories before their memory judgments than did participants in the few-memories condition [ $F(1,89) = 67,946, MS_e = 0.02, p < .001$ ]. There was no significant effect of the delay (1-month vs. 6-month sample) and no interaction ( $F_s < 1.8, n.s.$ ). Importantly, a manipulation check for recall difficulty (see Table 1) showed that participants who recalled a large number of memories rated the recall task as more difficult ( $M = 4.41, SD = 1.82$ ) than did participants who recalled only a small number of memories ( $M = 2.55, SD = 1.91$ ) [ $F(1,89) = 25.57, MS_e = 3.19, p < .001$ ]. Retrieval was perceived as more difficult after a 6-month delay than after a 1-month delay [ $F(1,89) = 9.85, p < .01$ ], but there was no significant interaction between the number of memories and the delay ( $F < 1, n.s.$ ).

**Recall performance.** The mean number of memories participants recalled before the memory judgment (many-memories condition) and both before and after the memory judgment (few-memories condition) was 11.96 ( $SD = 0.26$ ). Thus, virtually all participants fully complied with the recall task by reporting the required number of different memories. A  $2 \times 2$  ANOVA revealed that overall recall performance did not differ reliably either between

the few-memories condition ( $M = 11.94$ ) and the many-memories condition ( $M = 11.98$ ) or between the 1-month delay sample ( $M = 11.95$ ) and the 6-month delay sample ( $M = 11.96$ ) (all  $F_s < 1, n.s.$ ). Thus, eventual differences in memory judgments could not be due to unequal levels of overall memory performance. Before their memory judgments, all participants in the few-memories condition reported the requested number of memories (i.e., four).

**Judgments of memory quality.** Participants' ratings of the quality of their memories are shown in Table 1. Importantly, participants judged their memory as worse after retrieving many memories from the target episode ( $M = 6.77, SD = 1.33$ ) than after retrieving only a few memories ( $M = 7.71, SD = 1.54$ ) [ $F(1,89) = 9.77, MS_e = 2.11, p < .01$ ]. There was no significant main effect of delay and no significant number of memories  $\times$  delay interaction ( $F_s < 1$ ). Planned contrasts revealed a significant effect of number of recalled memories on memory judgments both after a 1-month delay [ $F(1,89) = 4.84, p < .05$ ] and after a 6-month delay [ $F(1,89) = 4.94, p < .05$ ].

**Mediational analyses.** We followed the procedure recommended by Baron and Kenny (1986) to test whether the effect of the number of recalled memories on memory judgments was mediated by the experienced ease of retrieval. We tested four criteria of mediation suggested by Baron and Kenny: (1) the independent variable (i.e., number of memories) significantly predicts the dependent variable (i.e., judgment of memory quality); (2) the independent variable significantly influences the proposed mediator (i.e., ease of retrieval); (3) the proposed mediator significantly predicts the dependent variable; and (4) when the proposed mediator and the independent variable are simultaneously entered in a regression analysis as predictors of the dependent variable, the influence of the proposed mediator remains significant, whereas the influence of the independent variable is significantly reduced. Ratings of recall difficulty were recoded so that higher values reflected higher ease (rather than higher difficulty). Thus, the naming of this variable was consistent with the conceptual framework.

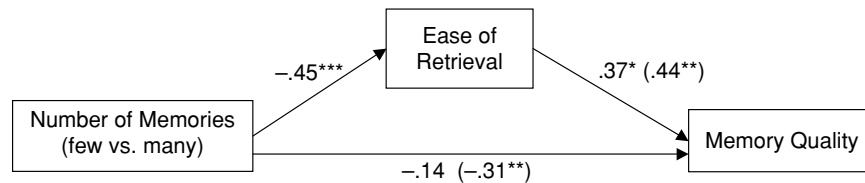
The results of the complete mediational analysis are presented in Figure 1. As expected from the ANOVAs above, regression analyses revealed that ratings of memory quality were lower after the recall of many specific memories than after the recall of few [ $\beta = -.31, t(91) = 3.14, p < .01$ ] (Criterion 1) and that retrieval ease was significantly

**Table 1**  
**Experiment 1: Mean Ratings of Recall Difficulty and Memory Quality as a Function of Number of Recalled Memories and Retention Interval**

	1-Month Interval		6-Month Interval	
	Few Memories	Many Memories	Few Memories	Many Memories
Difficulty of retrieval	1.83	3.90	3.19	4.87
Memory quality	7.87	6.90	7.58	6.65

Note—Judgments were made on a 9-point scale (higher values indicate higher ratings). In the few-memories/1-month condition,  $n = 23$ ; in the few-memories/6-month condition,  $n = 26$ ; in the many-memories/1-month condition,  $n = 21$ ; in the many-memories/6-month condition,  $n = 23$ .





**Figure 1. Mediation analysis for Experiment 1 with number of memories (few vs. many) as an independent variable, experienced ease of retrieval (recoded from ratings of difficulty) as a mediator, and rating of memory quality as the dependent variable. Path coefficients are standardized  $\beta$  coefficients from (multiple) regression analyses. The numbers in parentheses represent the direct effect (bivariate  $\beta$  coefficients) of one of the two predictors (i.e., number of memories or ease of retrieval) on ratings of memory quality prior to the inclusion of the other predictor. \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .**

related to the number of previously recalled memories [ $\beta = -.45$ ,  $t(91) = 4.80$ ,  $p < .001$ ] (Criterion 2). We also found that participants who experienced retrieval as easier exhibited higher ratings of memory quality [ $\beta = .44$ ,  $t(91) = 4.66$ ,  $p < .001$ ] (Criterion 3). Most importantly, when number of memories and ease of retrieval were included as predictors of memory judgments, only experienced ease predicted the judgment of memory quality significantly [ $\beta = .37$ ,  $t(90) = 3.56$ ,  $p < .001$ ], whereas the  $\beta$  coefficient for the number of memories was reduced to nonsignificance [ $\beta = -.14$ ,  $t(90) = 1.37$ ,  $p = .17$ ]. The indirect effect of number of memories on memory quality via ease of retrieval was significant in a Sobel test of mediation (Sobel, 1982;  $z = 2.87$ ,  $p < .01$ ) (Criterion 4). This pattern of findings suggests that the effect of the recalled number of memories on memory judgments was at least partially mediated by the effect of participants' experienced ease of retrieving their memories.

**Vividness as a content-based cue.** Ratings of vividness were significantly higher in the few-memories condition ( $M = 7.22$ ,  $SD = 1.55$ ) than in the many-memories condition ( $M = 6.16$ ,  $SD = 1.61$ ) [ $F(1,89) = 10.76$ ,  $MS_e = 2.52$ ,  $p < .01$ ]. Also, ratings of memory vividness were significantly correlated with judgments of memory quality, both in a bivariate regression ( $\beta = .64$ ,  $p < .001$ ) and, more importantly, in the multiple regression used in the mediational analysis ( $\beta = .54$ ,  $p < .001$ ). When vividness was included as an additional predictor in this regression analysis, the  $\beta$  coefficient for retrieval ease remained significant ( $\beta = .22$ ,  $p = .02$ ). It should be noted that the close relation between the vividness ratings and memory quality could be due to the sequence of the two ratings in the test booklet: Participants rated the vividness of their memories always shortly after they rated memory quality. Because the two ratings are likely to strike participants as related, they may have tried to exhibit a consistent response pattern in their ratings. Indeed, some of the participants reported in the postexperimental questioning seeing a connection between vividness and memory quality.

In sum, the analyses of Experiment 1 clearly demonstrate that participants recalling many specific memories of a single-event episode (the previous New Year's Eve) judged their memory for that event to be worse than did participants recalling only a few specific memories. Im-

portantly, both a manipulation check and a formal mediation analysis revealed that the effects on judgments of memory quality were mediated by the experienced ease of retrieving the specific memories. Notably, whereas vividness ratings explained a substantial share of variance in the memory judgments, the inclusion of vividness did not eliminate the significant contribution of experienced ease. Thus, although the contribution of vividness may have been overestimated because of the rating order (see above), ease of retrieval still had a nonredundant (i.e., an independent) effect on memory judgments. In other words, the presence of a diagnostic content-based cue did not forestall the use of the feeling-based cue.

One cannot conclude from these findings that ease of retrieval is taken into account when making judgments about all types of episodic memories. As we already noted in the introduction, so-called flashbulb memories may be one possible exception. An opportunity to study this issue arose shortly after the September 11 terrorist attacks. One of the authors of this article (W.H.) is based in New York City; the other author (G.E.) was working as a postdoctoral fellow in New York City at the time of the attacks, and shortly afterward returned to Germany. As we argued above, the use of ease of retrieval may depend on the character of the target event, specifically on the concurrent perception of the event as more or less shocking. Experiment 2 was conducted to explore this possibility by comparing memory judgment processes for a nonshocking event with those for a shocking event (learning about the September 11 attacks). The shocking nature of 9/11 is unquestionable (see Cohn, Mehl, & Pennebaker, 2004): Many people experienced negative emotions (Cohn et al., 2004) or even suffered from pronounced emotional distress (Galea et al., 2002; Silver, Holman, McIntosh, Poulin, & Gil-Rivas, 2002).

## EXPERIMENT 2

As in Experiment 1, participants were again asked to recall either few (4) or many (12) memories before their memory judgments. In one condition, participants reported memory judgments for the same nonshocking event as in Experiment 1 (i.e., the previous New Year's Eve), whereas those in two other conditions provided memory judg-

ments for a shocking event—that is, learning about the terrorist attacks on September 11, 2001. The critical difference between the two groups remembering their 9/11 experiences was the degree of shock present at the time of recall: Whereas the 9/11 events should still be perceived as highly shocking in temporal or physical proximity to the terrorist attacks, their shocking character may subside after a few months, especially outside the U.S. With these speculations in mind, we collected data from participants in Germany 1 month after the attacks (temporal proximity) and from participants in New York City 6 months after (physical proximity) and grouped them into what we will call the *high-shock condition*. We also collected data from participants in Germany 6 and 24 months after 9/11 (temporal and physical distance). This group will be referred to as the *attenuated-shock condition*.

## Method

**Participants.** A total of 322 students from the New School University in New York City and from the University of Cologne (Germany) participated in this study (256 female, 65 male, 1 transgender; mean age = 27.1 years). Out of these, 108 participants were included in the high-shock condition, which was run in Germany 1 month after 9/11 (50 participants) and in New York City 6 months after 9/11 (58 participants). A further 112 participants were included in the attenuated-shock condition, which was run in Germany 6 months (51 participants) and 24 months (61 participants) after 9/11. Another 102 participants were included in a no-shock condition (memories for the previous New Year's Eve), which was run after delays of 1 month (52 participants) and 6 months (50 participants) in Germany. The demographic variables did not differ significantly between the test conditions and were not significantly correlated with the dependent measures. The psychology students among the participants received curricular credit for their participation; all others received no compensation.

**Design.** The main study was based on a 2 (number of memories: few vs. many)  $\times$  3 (shock level: high shock vs. attenuated shock vs. no shock) between-participants design. The dependent variables and manipulation checks were the same as in Experiment 1, except that the ratings in the high-shock and the attenuated-shock conditions referred to memories for 9/11. In addition, a no-task condition was used to verify our claims about the emotional involvement of the participants without the confound of the experimental manipulation. The participants were randomly assigned to the few-memories, many-memories, or no-task condition.

**Materials and Procedure.** The materials and procedure were the same as in Experiment 1, except for the following modifications. The participants in the high-shock and attenuated-shock conditions received a test booklet titled "Memory for September 11." The recall task instructions closely followed the wording in Experiment 1, except that they referred to the event of learning about the terrorist attacks of 9/11. The participants were told not to include facts or pieces of information that they had learned from the media or the press. The instruction for judging the quality of one's 9/11 memory read: "How good is your memory for the circumstances under which you learned about or experienced the terrorist attacks on September 11?" (1 = *very poor*, 9 = *very good*).

In the no-task condition, the participants were not given a recall task before rating their event memory. These participants rated their emotional involvement concerning the target event on a 9-point scale (1 = *not involved at all*, 9 = *very involved*). The ratings of memory quality and vividness were then elicited in the same way as in the few- and many-memories conditions.

**Assessment of recall performance.** As in Experiment 1, we conducted a content analysis to determine the number of acceptable

responses for each participant. In the high-shock and attenuated-shock conditions, acceptable responses were those that referred to personal experiences during the target episode (learning about the 9/11 events). An acceptable response, for instance, was "A colleague told me to switch on the TV." Intercooder agreement was sufficiently high ( $\kappa = .80$ ). Five participants (2 in the high-shock condition, 2 in the attenuated-shock condition, and 1 in the no-shock condition) were excluded from the analyses because their recall reports contained a large number of unacceptable responses (from 6 to 10). For the sake of simplicity, these participants are not included among those described in the Participants section above.

## Results and Discussion

**Homogeneity within samples.** In order to determine the homogeneity of the samples subsumed within the different shock-level conditions, mean scores on the main dependent variables were compared. In each of the three samples (high-, attenuated-, and no-shock samples), a 2 (subsample)  $\times$  2 (number of memories) ANOVA was run for each of the four dependent measures. The four ANOVAs involving the high- and attenuated-shock samples did not reveal any significant main effects or interactions (all  $F$ s < 1.66, all  $p$ s > .20). As for the no-shock sample, ratings of memory quality, vividness, and emotional involvement were higher 1 month after the target event than 6 months after the event (all  $p$ s < .05), and ratings of difficulty were significantly lower ( $p < .001$ ) after the 1-month interval than after the 6-month interval. These findings are predicted by standard models of forgetting in episodic memory. More importantly for the present focus, there was no significant subsample  $\times$  number of memories interaction in the no-shock condition (all  $F$ s < 1.8, all  $p$ s > .20), suggesting that the effect of the recall task (few vs. many memories) did not differ across the two subsamples for the nonshocking event.

**Manipulation checks.** Overall, participants in the many-memories condition recalled significantly more memories before their memory judgments than did participants in the few-memories condition [ $F(1,209) = 973.2$ ,  $MS_e = 2.37$ ,  $p < .001$ ]. Thus, participants did follow the instructions for the recall task. In addition, a manipulation check for recall difficulty (see Table 2, top row) showed that participants who recalled a large number of memories rated the recall task as more difficult ( $M = 4.17$ ,  $SD = 2.02$ ) than did participants who recalled only a small number of memories ( $M = 2.63$ ,  $SD = 1.55$ ) [ $F(1,209) = 38.83$ ,  $MS_e = 3.21$ ,  $p < .001$ ]. Neither shock level nor the interaction between number of memories and shock level had significant effects [ $F(2,209) = 1.41$ ,  $p = .25$ , and  $F(1,209) = 1.56$ ,  $p = .21$ , respectively]. The differences in perceived difficulty were corroborated by tests of simple main effects within each of the three shock-level conditions [ $F$ s(1,209)  $\geq 5.97$ , all  $p$ s < .01].

A comparison of memory ratings across different shock levels in the no-task condition yielded the anticipated results (see Table 3). Clearly, ratings of emotional involvement, memory quality, and vividness differed across the three groups [ $F(2,103) = 12.87$ ,  $MS_e = 4.05$ ,  $p < .001$ ;  $F(2,104) = 5.29$ ,  $MS_e = 1.74$ ,  $p < .01$ ; and  $F(2,104) = 13.64$ ,  $MS_e = 2.10$ ,  $p < .001$ , for the different ratings, re-

**Table 2**  
**Experiment 2: Mean Ratings of Recall Difficulty, Memory Quality, and Vividness of Memories as a Function of Number of Recalled Memories and the Shocking Character of the Remembered Event**

	No Shock		Attenuated Shock		High Shock	
	Few Memories	Many Memories	Few Memories	Many Memories	Few Memories	Many Memories
Difficulty of retrieval	2.69	4.11	2.59	4.69	2.61	3.67
Memory quality	7.63	6.58	8.27	7.14	7.94	8.12
Vividness	7.34	6.47	7.78	6.94	7.78	7.39

Note—In the no-shock condition, participants recalled experiences on the previous New Year’s Eve. Participants in the high-shock and attenuated-shock conditions provided memory judgments for their experiences surrounding “September 11.” In the attenuated-shock condition, memory judgments were elicited from participants in Germany 6 and 24 months after 9/11 (temporal and physical distance). In the high-shock condition, memory judgments were provided by participants in Germany 1 month after the attacks (temporal proximity) and participants in New York City 6 months after 9/11 (physical proximity). Judgments were made on a 9-point scale (higher values indicate higher ratings); *n* was between 33 and 38 per cell.

spectively; the lower *df* for emotional involvement was due to a missing value]. Participants in the high-shock condition indicated significantly higher emotional involvement, memory quality, and vividness relative to those in either the no-shock [ $t(103) = 4.63, p < .001; t(104) = 2.89, p < .01$ ; and  $t(104) = 4.21, p < .001$ , for the three ratings, respectively] or the attenuated-shock [ $t(103) = 3.97, p < .001; t(104) = 2.66, p < .01$ ; and  $t(104) = 4.68, p < .001$ , respectively] condition. All other paired contrasts were nonsignificant ( $ts < 1, n.s.$ ).

**Recall performance.** Importantly, neither the main effect of number of memories nor the interaction between number of memories and shock level were significant ( $F_s < 1, n.s.$ ). The number of recalled memories did differ significantly across the different shock levels of the target event [ $F(2,209) = 14.75, MS_e = 5.01, p > .001$ ]. The mean number of recalled memories was higher in the high-shock condition ( $M = 10.90$ ) than in the attenuated-shock condition ( $M = 9.69$ ) [ $F(1,209) = 10.39, p < .01$ ], and these mean scores were both lower relative to the no-shock condition ( $M = 11.69$ ) [ $F(1,209) = 4.38, p < .05$ , and  $F(1,209) = 29.08, p < .001$ , respectively]. Thus, whereas most participants recalling their New Year’s Eve experiences reported the required number of memories, substantially fewer participants recalling their 9/11 experi-

ences listed the required number. Across all shock levels of the target event, participants in the few-memories condition reported the requested number of memories (i.e., four).

Participants who do not achieve the task criterion may feel that they perform worse compared to participants who fully reach the required number of memories. Lower perceived task accomplishment may, in turn, lead to lower judgments of memory quality. Thus, the reported ease-of-retrieval effects might have been driven by participants who did not recall all of the requested memories in the many-memories condition. In order to control for this possible confound, the main analyses were recalculated for a subset of participants who fully completed the recall task (see Belli et al., 1998). Even given the loss of statistical test power, the statistical inferences drawn from the restricted data set matched those drawn from the unrestricted data set.<sup>2</sup>

**ANOVA for judgments of memory quality.** Participants’ ratings of memory quality are shown in Table 2 (middle row). Importantly, the effect of the number of memories depended on the specific nature of the remembered event: An ANOVA revealed main effects of number of memories [ $F(1,209) = 11.49, p < .001$ ] and shock level [ $F(2,209) = 7.34, p < .001$ ], which were qualified by a significant interaction between the two factors [ $F(2,209) = 4.51, MS_e = 2.14, p < .05$ ]. Further analysis demonstrated that whereas memory judgments were significantly lower after the recall of many (vs. few) instances in the no-shock [ $F(1,209) = 9.37, MS_e = 2.14, p < .01$ ] and the attenuated-shock [ $F(1,209) = 11.45, p < .001$ ] conditions, there was no such effect when the shock of the target event was still strong ( $F < 1, n.s.$ ).

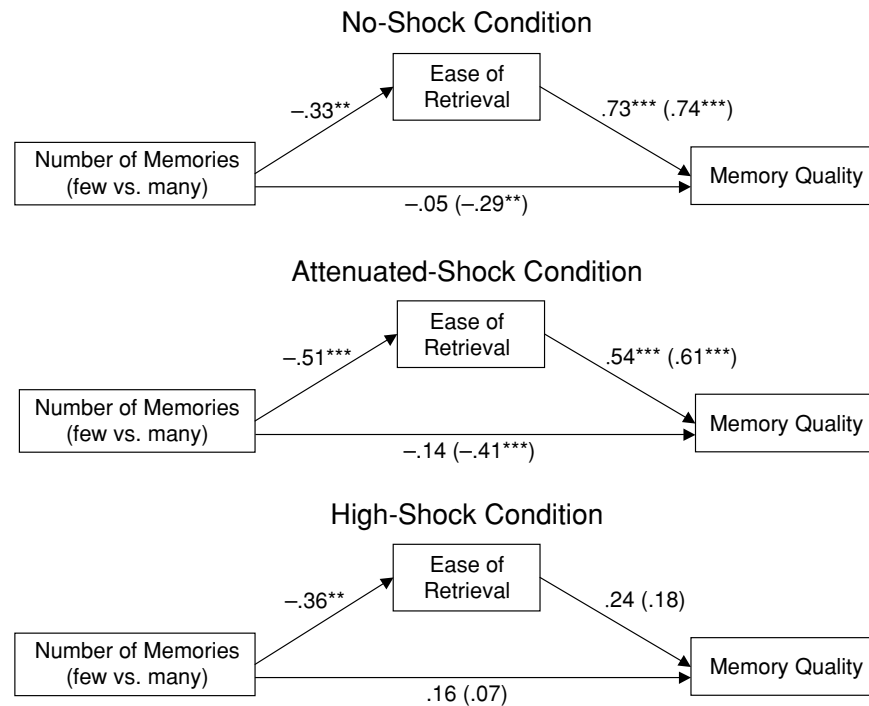
**Mediational analyses.** As in the data analysis for Experiment 1, we followed the procedure recommended by Baron and Kenny (1986) to investigate whether ease of retrieval mediated the number-of-memories effects on memory judgments in the attenuated-shock and no-shock conditions. The four criteria of mediation outlined by Baron and Kenny were tested in both conditions (see Experiment 1). Figure 2 summarizes the findings from these analyses.

Regression analyses for the no-shock and attenuated-shock conditions revealed that ratings of memory quality

**Table 3**  
**No-Task Condition in Experiment 2: Mean Ratings (and Corresponding Standard Deviations) of Emotional Involvement, Memory Quality, and Vividness of Memories as a Function of the Shocking Character of the Remembered Event**

	No Shock		Attenuated Shock		High Shock	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Emotional involvement	4.59	1.90	5.05	2.20	6.87	1.91
Memory quality	7.55	1.43	7.69	1.63	8.49	0.79
Vividness	6.66	1.37	6.62	1.83	8.15	1.01

Note—Participants in this control condition were not asked to recall specific memories before making their ratings. All judgments were made on 9-point scales (higher values indicate higher ratings). In the no-shock condition,  $n = 29$ ; in the attenuated-shock condition,  $n = 39$  ( $n = 38$  for emotional involvement due to a missing value); in the high-shock condition,  $n = 39$ .



**Figure 2. Experiment 2: Separate mediational analyses for memory judgments for different shock levels of the remembered event (no-shock, attenuated-shock, and high-shock conditions) with number of memories (few vs. many) as an independent variable, experienced ease of retrieval (recoded from ratings of difficulty) as a mediator, and rating of memory quality as the dependent variable. Path coefficients are standardized  $\beta$  coefficients from (multiple) regression analyses. The numbers in parentheses represent the direct effect (bivariate  $\beta$  coefficients) of one of the two predictors (i.e., number of memories or ease of retrieval) on ratings of memory quality prior to the inclusion of the other predictor. \*\* $p < .01$ . \*\*\* $p < .001$ .**

were lower after the recall of many rather than few specific memories [ $\beta = -.29$ ,  $t(71) = 2.50$ ,  $p < .01$ , and  $\beta = -.41$ ,  $t(71) = 3.83$ ,  $p < .001$ , respectively] (Criterion 1). Also, participants' experienced ease was significantly related to the number of previously recalled memories [ $\beta = -.33$ ,  $t(71) = 2.91$ ,  $p < .01$ , and  $\beta = -.51$ ,  $t(71) = 4.94$ ,  $p < .001$ , for the no-shock and attenuated-shock conditions, respectively] (Criterion 2). Furthermore, participants experiencing retrieval as easier exhibited higher ratings of memory quality both in the no-shock [ $\beta = .74$ ,  $t(71) = 9.34$ ,  $p < .001$ ] and attenuated-shock [ $\beta = .61$ ,  $t(71) = 6.53$ ,  $p < .001$ ] conditions (Criterion 3). Most importantly, when number of memories and ease of retrieval were simultaneously included as independent variables, only experienced ease predicted the judgment of memory quality significantly [ $\beta = .73$ ,  $t(70) = 8.60$ ,  $p < .001$ , and  $\beta = .54$ ,  $t(70) = 5.01$ ,  $p < .001$ , in the no-shock and attenuated-shock conditions, respectively], whereas the  $\beta$  coefficient for the number of memories was reduced to nonsignificance [ $\beta = -.05$ ,  $t(70) = 0.56$ ,  $p = .58$ , and  $\beta = -.14$ ,  $t(70) = 1.29$ ,  $p = .20$ , respectively]. The indirect effect of number of memories on memory quality via ease of retrieval was significant in Sobel (1982) tests of mediation for both types of target events [ $z = 2.76$ ,  $p < .01$ , and  $z = 3.51$ ,  $p < .001$ , for the no-shock and attenuated-shock conditions, respectively] (Criterion 4). Thus, the

effect of the number of recalled memories on memory judgments was at least partially mediated by the effect of participants' experienced ease of retrieval when the target event either was not shocking at all or had ceased being highly shocking.

In contrast, when the target event was highly shocking, neither number of recalled memories nor ease of retrieval was a significant predictor of judgments of memory quality [ $\beta = .07$ ,  $t(67) = 0.60$ ,  $p = .55$ , and  $\beta = .18$ ,  $t(67) = 1.50$ ,  $p = .14$ , respectively]. As expected from the significant manipulation check, number of memories did have a significant effect on ease of retrieval [ $\beta = -.36$ ,  $t(67) = 3.15$ ,  $p < .01$ ]. When number of memories and ease of retrieval were both included as independent variables, ease of retrieval had only a marginally significant effect on memory judgments [ $\beta = .24$ ,  $t(66) = 1.84$ ,  $p = .07$ ], and the effect of number of memories remained nonsignificant [ $\beta = .16$ ,  $t(66) = 1.23$ ,  $p = .22$ ]. Clearly, in the high-shock condition judgments of memory quality could not be predicted by either the simultaneous or the separate contributions of number of memories and retrieval ease.

Thus, when there was an effect of number of memories on memory judgments (as in the no-shock and attenuated-shock conditions), this effect was mediated by the experienced ease of retrieving specific memories from the target episode. When there was no such effect (as in the high-



shock condition), the contributions of both number of memories and retrieval ease in predicting memory judgments were rather low.

**Moderation by shock level.** The analyses in the previous section provided descriptive evidence that the strength of association between the independent variables (ease of retrieval and number of memories) and memory judgments differed across the three shock levels of the remembered event. In statistical terms, the relation between the independent variables and the dependent variable appeared to be moderated by shock level. This moderation will now be more formally examined by means of multiple regression, including interaction terms (see Aiken & West, 1991; Baron & Kenny, 1986), in order to allow statistical inferences about the strength of moderation, rather than mere description. Introducing interaction terms in a multiple regression enables us to test the hypothesis that the effect of one independent variable (here, either number of memories or retrieval ease) depends on another independent variable (the categorical variable shock level).

Contrast coding was used to represent the different shock levels and the number-of-memories conditions (see Cohen, Cohen, West, & Aiken, 2003). Shock level was contrast coded in order to compare the high-shock condition (coded as +2) with both the attenuated-shock (coded -1) and the no-shock (also coded -1) conditions. Thus, we could test whether the independent variables (ease and number of memories) were less closely associated with memory judgments when the remembered event was highly shocking than when it was less or not at all shocking. Number of memories was contrast coded in order to compare the many-memories condition (coded +1) with the few-memories condition (coded -1). In order to examine moderating influences of shock level on (1) the number-of-memories effect and (2) the retrieval ease effect, two interaction terms were computed—number of memories  $\times$  shock level and retrieval ease  $\times$  shock level.

The reported mediation of the number-of-memories effect through retrieval ease suggested a comparison of two regression models—that is, a restricted model including number-of-memories but not retrieval ease effects

(Model 1), and an extended one including terms for both effects (Model 2). The continuous predictor (retrieval ease) was centered by means of a  $z$  transformation (see Aiken & West, 1991). Neither higher order interactions (e.g., three-way interaction terms) nor the number of memories  $\times$  retrieval ease interaction are reported here, because they were not significant in additional regression analyses conducted during the preparation of this article. As can be seen in Table 4, the contrast-coded shock level significantly predicted memory judgments in both models, with a highly shocking event involving higher judgments of memory quality. Also, number of memories had a significant effect in Model 1 that was qualified by a significant number of memories  $\times$  shock level interaction. In contrast, when memory judgments were regressed on the extended set of predictors in Model 2, retrieval ease had a significant effect, qualified by a significant retrieval ease  $\times$  shock level interaction, whereas all effects involving number of memories were nonsignificant. These findings present strong evidence that (1) memory judgments depended on the ease of retrieving specific memories rather than on the number-of-memories manipulation, and (2) the ease-of-retrieval effect was significantly moderated by the shocking character of the target event. As expected from the previous analyses, the signs of the  $\beta$  coefficients indicated that memory judgments depended less on retrieval ease when the remembered event was highly shocking than when it was less or not at all shocking. It should be noted that Model 2 explained substantially more variance (adjusted  $R^2 = .37$ ) than did Model 1 (adjusted  $R^2 = .11$ ), testifying to the important role of retrieval ease and its moderation by shock level in predicting memory judgments.<sup>3</sup>

**Vividness as a content-based cue.** Ratings (see Table 2, bottom row) were higher for 9/11 memories than for New Year's Eve memories [ $F(1,209) = 5.74, p < .01$ , and  $F(1,209) = 2.68, p = .051$ , for the comparisons of the high- and attenuated-shock conditions, respectively, with the no-shock condition]. Ratings of vividness were significantly lower after the recall of many rather than few memories only for the no-shock [ $F(1,209) = 4.86, p < .05$ ] and attenuated-shock [ $F(1,209) = 4.54, p < .05$ ]

**Table 4**  
Standardized Regression Coefficients ( $\beta$ ) and Corresponding  $t$  Statistics  
From Different Multiple Regression Analyses of Memory Judgments

Predictor	Model 1		Model 2	
	$\beta$	$t$	$\beta$	$t$
Number of memories (many vs. few)	-.22	-3.41***	-.01	-.23
Retrieval ease	—	—	.49	7.78***
Shock level (high vs. attenuated/no shock)	.19	2.94**	.16	2.86**
Number of memories $\times$ shock level	.19	3.00**	.09	1.62
Retrieval ease $\times$ shock level	—	—	-.15	-2.45**

Note—These analyses were conducted across all three shock-level conditions. Shock level was contrast coded to compare the high-shock condition (coded as +2) with both the attenuated-shock condition (coded as -1) and the no-shock condition (also coded -1). Number of memories was contrast coded to compare the many-memories condition (coded as +1) with the few-memories condition (coded as -1). Adjusted  $R^2 = .11$  and .37 for Models 1 and 2, respectively;  $df = 214$  for all  $t$  statistics. Retrieval ease was based on recoded ratings of retrieval difficulty, with higher scores reflecting greater ease. See the text for further explanations. \*\* $p < .01$ . \*\*\* $p < .001$  (one-tailed).

conditions, not for the high-shock condition ( $F < 1$ , n.s.). The number of memories  $\times$  shock level interaction was not significant ( $F < 1$ ). When vividness was included as an additional predictor in the previous regression models, it had a significant effect, which did not eliminate the other, previously found effects. Across all shock levels, vividness was significantly correlated with memory judgments, both in a bivariate regression ( $\beta = .68, p < .001$ ) and in the regression model used in the mediational analysis (i.e., a model with number of memories and retrieval ease as predictors;  $\beta = .52, p < .001$ ). When vividness was included as an additional predictor in this model, the  $\beta$  coefficient for experienced ease remained highly significant ( $\beta = .30, p < .001$ ). In three separate regression analyses for each of the different shock levels, vividness was a significant predictor (all  $\beta$ s between .43 and .50,  $ps < .001$ ), and retrieval ease remained significant in the attenuated-shock and no-shock conditions ( $\beta = .24, p < .01$ , and  $\beta = .48, p < .001$ , respectively). Furthermore, when vividness was included in the regression models testing moderation, it had a highly significant effect ( $\beta$ s  $> .47, ps < .001$ ), yet all significant effects reported for Models 1 and 2 still remained significant, with approximately the same  $\beta$  coefficients and  $t$  values as in the analyses above (see Table 4). The effects found for retrieval ease as well as the moderation of retrieval ease effects by shock level were independent of the contribution of vividness.

In sum, the mechanisms underlying judgments of memory quality differed depending on the perceived status of the event at the time of remembering. Findings from a no-task control condition confirmed the assumption that participants in close temporal or physical proximity to the 9/11 events perceived the target event as highly shocking: Consistent with the literature on flashbulb memories (e.g., Brown & Kulik, 1977; Christianson, 1992; Talarico & Rubin, 2003), the highly shocking public event was remembered with greater emotional involvement, greater confidence in the quality of one's memory, and higher vividness than was reported in the other conditions. Across all shock levels of the target event, the recall of many memories was accompanied by lower ratings of retrieval ease, but only in the attenuated-shock and no-shock conditions were these feeling-based differences reflected by rememberers' memory judgments. In fact, when 9/11 experiences were perceived as highly shocking, memory judgments were even slightly, although nonsignificantly, higher after the recall of many than after the recall of few specific memories. In contrast, when the same 9/11 experiences seemed more remote and less shocking (attenuated shock) or when the remembered event was not shocking, memory judgments were significantly lower after the recall of many rather than of few memories. This effect of number of memories was mediated by the experienced ease of retrieving these memories, a finding that replicates a main result of Experiment 1. Additional regression analyses formally confirmed the descriptive finding that the ease-of-retrieval effect was moderated by the shocking character of the remembered event. Because the target event was exactly the same in the high-shock and attenuated-

shock conditions, differences other than the current shocking character of the remembered event can hardly account for the pattern of findings. Correlational evidence suggests that our participants also took into account their memories' vividness (i.e., a content-based cue) to derive judgments of memory quality.

## GENERAL DISCUSSION

Inferences, heuristics, and educated guesses play a significant role in memory (see, e.g., Bransford & Johnson, 1972; Koriat, Goldsmith, & Pansky, 2000). For instance, they may affect remember-know judgments (Strack & Förster, 1995), judgments about the source of memory (e.g., Mather, Johnson, & DeLeonardis, 1999), judgments of learning (e.g., Benjamin et al., 1998), or the feeling of knowing (e.g., Koriat, 1993). The present research supports the notion that inferential processes also play a role in judgments about memory quality (see Skurnik et al., 2000). Several studies have shown that the experienced ease of retrieval can guide judgments about future memory performance (e.g., Begg et al., 1989; Benjamin et al., 1998) or about the quality of memory for a life period (Belli et al., 1998; Winkielman et al., 1998). Our findings demonstrate that ease of retrieval also guides judgments about the quality of memories for discrete episodes from the past. Hence, theorizing about how people judge their event memory during retrieval should take into account feeling-based cues in addition to well-researched content-based cues such as vividness or sensory detail (see, e.g., Conway, 1996; Conway & Pleydell-Pearce, 2000; Johnson et al., 1993; Johnson & Raye, 1981).

We manipulated ease of retrieval by using a technique that required participants to generate either many (12) or few (4) memories for a particular past episode, an adaptation of the methodology of Winkielman et al. (1998) and Belli et al. (1998). Our participants were more likely to judge their memory for an episode as harder to retrieve if they had previously generated many specific memories about the episode rather than only a few. The mediational analyses established that the effect of this manipulation on quality judgments was indirect, mediated through retrieval ease. This evidence essentially justifies using the term *ease-of-retrieval effect* on memory judgments, rather than referring to the principal finding as a *number-of-memories effect*.

Our findings also suggest that content-based cues such as vividness may play a yet underestimated role in demonstrations of ease-of-retrieval effects, although the close proximity of the memory quality and vividness ratings qualifies any definite conclusion (for the role of vivid imagery in memory, see, e.g., Brewer, 1986; Conway & Pleydell-Pearce, 2000; Garry et al., 1996). When we found ease-of-retrieval effects, manipulating number of memories not only affected experienced ease but also ratings of vividness (see Experiment 1 and the attenuated-shock and no-shock conditions in Experiment 2). Specifically, memories recalled in the many-memories condition were less vivid than those recalled in the few-memories

condition. This finding is consistent with the idea that memories provided toward the end of the recall list in the many-memories condition suffer in content attributes relative to memories provided earlier on in the list. Because of a recency effect, the less vivid (thus, relatively impoverished) memories at the end of the list may even be particularly salient at the time of making a memory judgment. Hence, previously reported effects of the number of recalled memories on memory judgments (Belli et al., 1998; Winkielman et al., 1998) may have resulted from differences in content attributes rather than from differences in retrieval ease per se. Going beyond the scope of these studies, the present research examined the roles of both vividness and retrieval ease in regression analyses. As a result, we found that although vividness was a highly significant predictor, there still was a nonredundant effect of retrieval ease on memory judgments. Thus, even if differences in memory judgments can be explained by differences in vividness, one must still reckon with ease of retrieval as an independent cue to memory judgments.

In addition to finding evidence that both content-based and feeling-based cues guided judgments of memory quality, we found that when participants were in shock about the attacks of September 11, 2001, they did not rely on retrieval ease when judging their corresponding episodic memories (Experiment 2). There are two possibilities as to how rememberers arrived at their memory judgments for such a highly shocking event. First, ratings of memory vividness were higher when the remembered event was highly shocking than when it was experienced as less shocking (as a result of temporal and physical distance from September 11) or when the event did not possess a shocking character in the first place (New Year's Eve). Also, participants in the high-shock condition did not rate their memories as less vivid when they had previously recalled a large number of memories than when they had recalled only a few. Thus, memories for the highly shocking event apparently had a degree of vividness that remained greatly unaffected by the manipulation of number of memories. This high and robust level of vividness may have been sufficiently powerful to affect memory judgments within the apparently small margin at the top of the rating scale. Thus, above a certain threshold level, vividness may be so salient that it forestalls the use of other cues, such as retrieval ease.

Second, the quality judgments could have been based on norms about what one ought to remember about shocking events. Refrains such as "I will never forget" were often heard, especially in the U.S., after the attacks of September 11 (for instance, see Web sites such as [www.wnfnsept11.com](http://www.wnfnsept11.com)). Such statements are ethical claims about a duty to remember (Margalit, 2002), and they also imply normative beliefs about what and how people ought to remember (see Jost, Kruglanski, & Nelson, 1998). In this vein, Costermans, Lories, and Ansay (1992) found that judgments about semantic memory were significantly correlated with the strength of normative beliefs, such as "I ought to know the answer because everyone certainly

knows it." Similarly, the participants in the high-shock condition may have reasoned: "Everyone ought to remember September 11 as if it were yesterday" or "I ought to remember how I learned about the attack because everyone else does." Such reasoning may have overshadowed participants' actual retrieval experiences. Thus, people may have judged episodic memories of their encounter with 9/11 on the basis of nonepisodic knowledge.

Further research is needed to investigate whether high levels of vividness or metamemory claims, or a combination of the two, are responsible for the observed differences between memory judgments for highly shocking versus less or nonshocking events. Clearly, effective experimental manipulations of both content-based cues and normative beliefs would be helpful in this respect. Future studies, then, may build on the main message of the present research: Feeling-based cues such as ease of retrieval can guide judgments for single events, yet when people are highly shocked about an event, they tend to place less emphasis on ease of retrieval.

## REFERENCES

- AIKEN, L. S., & WEST, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- ANDERSON, S. J., & CONWAY, M. A. (1993). Investigating the structure of autobiographical memories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 1178-1196.
- ARKES, H. R., HACKETT, C., & BOEHM, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making*, *2*, 81-94.
- BARON, R. M., & KENNY, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality & Social Psychology*, *51*, 1173-1182.
- BEGG, I., DUFT, S., LALONDE, P., MELNICK, R., & SANVITO, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory & Language*, *28*, 610-632.
- BELLI, R. F., WINKIELMAN, P., READ, J. D., SCHWARZ, N., & LYNN, S. J. (1998). Recalling more childhood events leads to judgments of poorer memory: Implications for the recovered/false memory debate. *Psychonomic Bulletin & Review*, *5*, 318-323.
- BENJAMIN, A. S., & BJORK, R. A. (1996). Retrieval fluency as a metacognitive index. In L. M. Reder (Ed.), *Implicit memory and metacognition* (pp. 309-338). Mahwah, NJ: Erlbaum.
- BENJAMIN, A. S., BJORK, R. A., & SCHWARTZ, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, *127*, 55-68.
- BLESS, H., STRACK, F., & WALTHER, E. (2001). Memory as a target of social influence? Memory distortions as a function of social influence and metacognitive knowledge. In J. P. Forgas & K. D. Williams (Eds.), *Social influence: Direct and indirect processes* (pp. 167-183). Philadelphia: Psychology Press.
- BRANSFORD, J. D., & JOHNSON, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning & Verbal Behavior*, *11*, 717-726.
- BREWER, W. F. (1986). What is autobiographical memory? In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 25-49). Cambridge: Cambridge University Press.
- BROWN, R., & KULIK, J. (1977). Flashbulb memories. *Cognition*, *5*, 73-99.
- CHRISTIANSON, S.-Å. (1992). Do flashbulb memories differ from other types of emotional memories? In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of "flashbulb" memories* (pp. 191-211). Cambridge: Cambridge University Press.
- COHEN, J., COHEN, P., WEST, S. G., & AIKEN, L. S. (2003). *Applied mul-*



- multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.
- COHN, M. A., MEHL, M. R., & PENNEBAKER, J. W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science*, **15**, 687-693.
- CONWAY, M. A. (1990). *Autobiographical memory: An introduction*. Buckingham, U.K.: Open University Press.
- CONWAY, M. A. (1996). Autobiographical knowledge and autobiographical memories. In D. C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 67-93). Cambridge: Cambridge University Press.
- CONWAY, M. A., & PLEYDELL-PEARCE, C. W. (2000). The constructions of autobiographical memories in the self memory system. *Psychological Review*, **107**, 261-288.
- COSTERMANS, J., LORIES, G., & ANSAY, C. (1992). Confidence level and feeling of knowing in question answering: The weight of inferential processes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 142-150.
- DEWHURST, S. A., & CONWAY, M. A. (1994). Pictures, images and recollective experience. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1088-1098.
- GALEA, S., AHERN, J., RESNICK, H., KILPATRICK, D., BUCUVALAS, M., GOLD, J., & VLAHOV, D. (2002). Psychological sequelae of the September 11 terrorist attacks in New York City. *New England Journal of Medicine*, **346**, 982-987.
- GARRY, M., MANNING, C. G., LOFTUS, E. F., & SHERMAN, S. J. (1996). Imagination inflation: Imagining a childhood event inflates confidence that it occurred. *Psychonomic Bulletin & Review*, **3**, 208-214.
- HADDOCK, G., ROTHMAN, A. J., REBER, R., & SCHWARZ, N. (1999). Forming judgments of attitude certainty, intensity, and importance: The role of subjective experiences. *Personality & Social Psychology Bulletin*, **25**, 771-782.
- HART, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, **56**, 208-216.
- HASHER, L., GOLDSTEIN, D., & TOPPINO, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning & Verbal Behavior*, **16**, 107-112.
- HERRMANN, D. J. (1982). Know thy memory: The use of questionnaires to assess and study memory. *Psychological Bulletin*, **92**, 434-452.
- JOHNSON, M. K., FOLEY, M. A., SUENGAS, A. G., & RAYE, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General*, **117**, 171-176.
- JOHNSON, M. K., HASHTROUDI, S., & LINDSAY, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3-28.
- JOHNSON, M. K., & RAYE, C. L. (1981). Reality monitoring. *Psychological Review*, **88**, 67-85.
- JOST, J. T., KRUGLANSKI, A. W., & NELSON, T. O. (1998). Social metacognition: An expansionist review. *Personality & Social Psychology Review*, **2**, 137-154.
- KELLEY, C. M., & LINDSAY, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory & Language*, **32**, 1-24.
- KORIAT, A. (1993). How do we know what we know? The accessibility account of the feeling of knowing. *Psychological Review*, **100**, 609-639.
- KORIAT, A., & GOLDSMITH, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, **103**, 490-517.
- KORIAT, A., GOLDSMITH, M., & PANSKY, A. (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology*, **51**, 481-537.
- KORIAT, A., & LEVY-SADOT, R. (1999). Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one's own knowledge. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 483-502). New York: Guilford.
- MARGALIT, A. (2002). *The ethics of memory*. Cambridge, MA: Harvard University Press.
- MATHER, M., JOHNSON, M. K., & DELEONARDIS, D. M. (1999). Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates. *Cognitive Neuropsychology*, **16**, 437-458.
- MAZZONI, G., & NELSON, T. O. (1995). Judgments of learning are affected by the kind of encoding in ways that cannot be attributed to the level of recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 1263-1274.
- MCCLOSKEY, M., WIBLE, C. G., & COHEN, N. J. (1988). Is there a special flashbulb-memory mechanism? *Journal of Experimental Psychology: General*, **117**, 171-181.
- METCALFE, J. (2000). Metamemory: Theory and data. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 197-211). Oxford: Oxford University Press.
- NEISSER, U., & HARSCH, N. (1992). Phantom flashbulb: 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). Cambridge: Cambridge University Press.
- NELSON, T. O. (1992). *Metacognition: Core readings*. Boston: Allyn & Bacon.
- NELSON, T. O., LEONESIO, R. J., SHIMAMURA, A. P., LANDWEHR, R. F., & NARENS, L. (1982). Overlearning and the feeling of knowing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **8**, 279-288.
- REISER, B. J., BLACK, J. B., & ABELSON, R. P. (1985). Knowledge structures in the organization and retrieval of autobiographical memories. *Cognitive Psychology*, **17**, 89-137.
- REISER, B. J., BLACK, J. B., & KALAMARIDES, P. (1986). Strategic memory search processes. In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 100-121). Cambridge: Cambridge University Press.
- ROSS, M., BUEHLER, R., & KARR, J. W. (1998). Assessing the accuracy of conflicting autobiographical memories. *Memory & Cognition*, **26**, 1233-1244.
- ROSS, M., & MACDONALD, T. K. (1997). How can we be sure? Using truth criteria to validate memories. In M. S. Myslobodsky (Ed.), *The mythomaniacs: The nature of deception and self-deception* (pp. 181-201). Mahwah, NJ: Erlbaum.
- ROTHMAN, A. J., & SCHWARZ, N. (1998). Constructing perceptions of vulnerability: Personal relevance and the use of experiential information in health judgments. *Personality & Social Psychology Bulletin*, **24**, 1053-1064.
- RUBIN, D. C. (2005). A basic-systems approach to autobiographical memory. *Current Directions in Psychological Science*, **14**, 79-83.
- RUBIN, D. C., & KOZIN, M. (1984). Vivid memories. *Cognition*, **16**, 81-95.
- SCHMIDT, S. R. (2004). Autobiographical memories for the September 11th attacks: Reconstructive errors and emotional impairment of memory. *Memory & Cognition*, **32**, 443-454.
- SCHWARZ, N., BLESS, H., STRACK, F., KLUMPP, G., RITTENAUER-SCHATKA, H., & SIMONS, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality & Social Psychology*, **61**, 195-202.
- SCHWARZ, N., & VAUGHN, L. A. (2002). The availability heuristic revisited: Ease of recall and content of recall as distinct sources of information. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 103-119). Cambridge: Cambridge University Press.
- SILVER, R. C., HOLMAN, E. A., MCINTOSH, D. N., POULIN, M., & GIL-RIVAS, V. (2002). Nationwide longitudinal study of psychological responses to September 11. *Journal of the American Medical Association*, **288**, 1235-1244.
- SKURNIK, I., SCHWARZ, N., & WINKELMAN, P. (2000). Drawing inferences from feelings: The role of naive beliefs. In H. Bless & J. P. Forgas (Eds.), *The message within: The role of subjective experience in social cognition and behavior* (pp. 162-175). Philadelphia: Psychology Press.
- SOBEL, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural models. In S. Leinhardt (Ed.), *Sociological methodology 1982* (pp. 290-312). San Francisco: Jossey-Bass.
- STRACK, F., & FÖRSTER, J. (1995). Reporting recollective experiences: Direct access to memory systems? *Psychological Science*, **6**, 352-358.
- TALARICO, J. M., & RUBIN, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science*, **14**, 455-461.
- TVERSKY, A., & KAHNEMAN, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, **5**, 207-232.
- WÄNKE, M., BOHNER, G., & JURKOWITSCH, A. (1997). There are many reasons to drive a BMW: Does imagined ease of argument generation influence attitudes? *Journal of Consumer Research*, **24**, 170-177.



- WEAVER, C. A. (1993). Do you need a "flash" to form a flashbulb memory? *Journal of Experimental Psychology: General*, **122**, 39-46.
- WHITTLESEA, B. W. A., & LEBOE, J. P. (2000). The heuristic basis of remembering and classification: Fluency, generation, and resemblance. *Journal of Experimental Psychology: General*, **129**, 84-106.
- WILLIAMS, J. M. G., HEALY, H. G., & ELLIS, N. C. (1999). The effect of imageability and predictability of cues in autobiographical memory. *Quarterly Journal of Experimental Psychology*, **52A**, 555-579.
- WINKIELMAN, P., SCHWARZ, N., & BELLI, R. F. (1998). The role of ease of retrieval and attribution in memory judgments: Judging your memory as worse despite recalling more events. *Psychological Science*, **9**, 124-126.
- WRIGHT, D. B. (1993). Recall of the Hillsborough disaster over time: Systematic biases of "flashbulb" memories. *Applied Cognitive Psychology*, **7**, 129-138.

#### NOTES

1. We thank Bob Belli for pointing out this aspect in response to an earlier version of this article.
2. Restricting the analysis to completers reduced case numbers in the many-memories condition from 33 to 23 in the high-shock condition,

from 36 to 14 in the attenuated-shock condition, and from 38 to 32 in the no-shock condition. However, the pattern of findings for the main analyses remained unchanged for the subset of completers. Whereas memory judgments were significantly lower after the recall of many (vs. few) instances in the no-shock condition [ $F(1,171) = 5.97$ ,  $MS_e = 2.16$ ,  $p < .01$ ] and the attenuated-shock condition [ $F(1,171) = 3.92$ ,  $p < .05$ ], there was no such effect when the target event was still highly shocking ( $F < 1$ , n.s.). Also, participants who recalled a large number of memories rated the recall task as more difficult than did participants who recalled only a small number of memories [ $F(1,171) = 25.36$ ,  $MS_e = 3.10$ ,  $p < .001$ ]. Neither shock level nor the interaction between number of memories and shock level had significant effects ( $F_s < 1$ , n.s.). Again, the differences in perceived difficulty were corroborated by tests of simple main effects within each of the three shock-level conditions [ $F_s(1,171) \geq 6.21$ ,  $ps < .01$ ].

3. It should be pointed out that the main pattern of findings (in particular, the statistical inferences) remained unchanged when the actual number of recalled memories was substituted for the contrast-coded independent variable (number of memories) in the regression analyses.

(Manuscript received August 19, 2003;  
revision accepted for publication May 18, 2005.)