

The Cognitive, Emotional, and Social Impacts of the September 11 Attacks: Group Differences in Memory for the Reception Context and the Determinants of Flashbulb Memory

OLIVIER LUMINET

*University of Louvain at Louvain-la-Neuve, Belgium
Belgian National Fund for Scientific Research*

ANTONIETTA CURCI

University of Bari, Italy

ELIZABETH J. MARSH

Washington University, St Louis

INEKE WESSEL

Maastricht University, The Netherlands

TICU CONSTANTIN

University of Iasi, Romania

FARUK GENCOZ

Middle East Technical University, Ankara, Turkey

MASAO YOGO

Doshisha University, Kyoto, Japan

ABSTRACT. The authors examined group differences in memories for hearing the news of and reactions to the September 11 attacks in 2001. They measured memory for reception context (immediate memory for the circumstances in which people first heard the news) and 11 predictors of the consistency of memory for reception context over time (flashbulb memory). Shortly after 9/11, a questionnaire was distributed to 3,665 participants in 9 countries. U.S. vs. non-U.S. respondents showed large differences in self-rated importance of the news and in memory for event-related facts. The groups showed moderate differences in background knowledge and emotional-feeling states. Within non-U.S. groups, there were large differences for emotional-feeling states and moderate differences for personal rehearsal, background knowledge, and attitudes toward the United States. The authors discuss the implications of those findings for the study of group differences in memory and for the formation of flashbulb memories.

Key words: emotions, group comparisons, memory, September 11

FOR MORE THAN TWO DECADES, emotion and memory researchers have been interested in people's memories for the circumstances in which they first hear about major public events, such as political assassinations (e.g., Brown & Kulik, 1977; Christianson, 1989) or natural or man-made disasters (e.g., Christianson & Engelberg, 1999; Neisser & Harsch, 1992; Wright, 1993). People report remembering seemingly mundane details such as their location, how they learned the news, what they were wearing, what they were doing, and with whom they first heard the news. Those memories, dubbed flashbulb memories by Brown and Kulik, have been the object of much research, including the examination of the amount of time required to observe distortion in those confidently held memories (Hornstein, Brown, & Mulligan, 2003; Schmolck, Buffalo, & Squire, 2000). In the present article, we focus on more immediate memories of hearing the news, which we call memories for the reception context. For a recollection to be considered a flashbulb memory, it must involve not only a live quality accompanied by recall of minutiae, but also preserve details of the reception events and remain unchanged over long periods of time (Conway, 1995). Acceptance of the last two conditions means that it is not technically possible to assess flashbulb memories shortly after the information has been encoded because a delayed retest is required to assess the preservation and consistency of details over time. *Memory for the reception context* is thus defined as people's immediate memory for the circumstances in which they first heard some important news. Only some of those memories will remain consistent, whereas others will fade or become distorted. Consistent memories for the reception context will be referred to as *flashbulb memories* in this article.

Across a number of different events, many people report vivid memories of hearing the news shortly after unexpected events. For example, 1 month after 96 soccer fans were crushed to death in the Hillsborough disaster in England, all of the respondents in a study by Wright (1993) reported memories for how they heard the news. Three days after the O. J. Simpson verdict was announced, more than 98% of respondents in a study by Schmolck et al. (2000) reported memories for reception context. Forty-eight hours after the *MS Estonia* ferry sank, 99% of participants in a study by Christianson and Engelberg (1999) reported memories for hearing the news. It is not clear, however, that those rates would have been as high with different participant populations.¹ Americans were unfamiliar with the Hillsborough disaster and the sinking of the *MS Estonia*. Europeans were much less interested in the O. J. Simpson trial than were Americans. Accordingly, there are studies that show the occurrence of more flashbulb memories in one group of participants than in another (e.g., Brown & Kulik, 1977; Conway et al., 1994; Curci, Luminet, Finkenauer, & Gisle, 2001; Neisser et al., 1996). For example, Brown and Kulik showed that, compared with Caucasian respondents, African American participants reported more memories for hearing the news of the assassinations of leaders in the U.S. civil rights movement (e.g., Malcolm X, Martin Luther King Jr.). Conway et al. looked at flashbulb memories for the

resignation of the British prime minister, Margaret Thatcher. A year after she resigned, more than 80% of British participants were classified as having flashbulb memories (i.e., vivid and consistent), whereas less than 30% of non-British participants were classified that way. Likewise, using the same criteria of vividness and consistency, Curci et al. found that French participants in their study had a significantly higher proportion of flashbulb memories for hearing about the death of the former French president, Francois Mitterrand, than did Belgian participants. Thus, the results of those comparison studies suggest that the salience of a major news event and its personal importance may vary for different cultural or national groups, which results in differences in flashbulb memories. One limitation of those studies, however, is that the groups that are typically compared are quite similar to one another (e.g., both Western countries). In addition, it is rare that more than two or three groups are compared.

As described heretofore, the results of a number of studies have shown group differences in flashbulb memories. In some cases, that means that years after the event, participants in one group are more likely to report remembering how they heard the news (e.g., Brown & Kulik, 1977). In other studies, it means that participants in one group have more consistent memories over time (e.g., Conway et al., 1994). However, although immediate memories are often collected and, indeed, are required to examine consistency over time, the data are rarely reported. We believe that is a mistake, because immediate memories are important for understanding the mechanisms underlying the eventual formation of flashbulb memories. For example, if participants do not initially have vivid memories for the reception context at Time 1, it would be unreasonable to expect them to have them at Time 2.

All authors are at the Department of Psychology of their respective universities.

This study was conducted by members of the Flashbulb Memory Network, an ongoing collaborative project on autobiographical memory. Parts of this study were supported by the Belgian National Fund for Scientific Research (Grant No. 1.5.146.02), the National Institute of Mental Health (1F32-MH12567), the Dutch Council for Scientific Research (NWO; 575-29-001); and the Japanese Society for the Promotion of Science (JSPS; 14710103).

The authors thank the following people for their contribution to various aspects of the study: Guglielmo Bellelli, Liesbeth van Breda, Giovanna D'Onchia, Dario Galati, Jenny van den Heuvel, Giannangela Mastrorilli, Prisca Niel, Annarita Pugliese, Mario Refice, Maud Slaats, Catherine Demaret, Jurgen Engerisser, Sylvie Grumiau, Véronique Pellegrin, Sophie Wautier, James Pennebaker, Allen Omoto, Martin Conway, Catrin Finke-nauer, Mark Snyder, Carolin Showers, Martial van der Linden, Marie-Claire Gay, Emmanuelle Zech, Anna Tcherkassof, Kate Niederhoffer, Ambroise Nève, Paula Niedenthal, Gwenola Herbette, Debra Golden, Gerard Duveen, and Annette Lafontaine.

Elizabeth J. Marsh is now in the Department of Psychological and Brain Sciences at Duke University.

Address correspondence to Olivier Luminet, University of Louvain at Louvain-la-Neuve, Department of Psychology, Research Unit for Emotion, Cognition, and Health, 10, Place du Cardinal Mercier, B-1348 Louvain-la-Neuve, Belgium; Olivier.Luminet@psp.ucl.ac.be (e-mail).

This article presents an empirical snapshot of the world's reaction to the 9/11 terrorist attacks. On September 11, 2001, more than 3,000 people died in the worst terrorist attacks ever committed on American soil. In a coordinated plan of action, four commercial airliners were hijacked and turned into missiles aimed at buildings that symbolized American prosperity. Two planes hit the north and south towers of the World Trade Center (WTC) in New York, leading to the eventual collapse of both buildings. A third plane hit the Pentagon near Washington, DC, and a fourth plane crashed in rural Pennsylvania. The attacks were allegedly planned and executed by members of Al Qaeda, a terrorist group based in Afghanistan.² Relatively soon after the attacks, we were able to obtain a large sample of people from nine different countries. Of interest to us was the effect of the differences in the countries' cultural and economic backgrounds on memory for reception context and on variables that are predictors of the later formation of flashbulb memories. Those variables included (a) novelty, (b) surprise, (c) personal consequentiality and importance, (d) valenced emotional-feeling states, (e) rehearsal, (f) background knowledge, (g) attitudes toward the target country, and (h) memory for event-related facts. We assumed that for Americans, the attacks of 9/11 would yield high ratings on all those variables (except for attitudes, which were not relevant for the U.S. group). We did not assume, however, that those rates would be as high as they were in other countries. The simultaneous testing of large groups of U.S. and non-U.S. respondents allowed us to examine possible group differences in those variables.

Before turning to our predictions, we will briefly review the literature on the variables that were measured in addition to memory for reception context. The dimensions of novelty and importance-consequentiality derive from cognitive theories of emotion (e.g., Frijda, 1986; Lazarus, 1982; Leventhal, 1984; Scherer, 1984). In those theories, emotions result from the appraisal of stimulus conditions in terms of their relevance for the organism and its behavioral adaptation. For the novelty-encoding hypothesis (Tulving & Kroll, 1995), novelty is a necessary condition for encoding information in long-term memory. The appraisal of novelty is then followed by a nonvalenced emotional state of surprise. The appraisal of importance-consequentiality develops almost simultaneously with the appraisal of novelty and has been proposed as a critical predictor for memory for the reception context (Brown & Kulik, 1977; Christianson, 1989; Christianson & Engelberg, 1999; Conway et al., 1994; Finkenauer et al., 1998; Rubin & Kozin, 1984). The appraisal of importance-consequentiality then leads to emotional-feeling states such as sadness, anger, or fear (Conway et al.; Finkenauer et al.; Oatley & Johnson-Laird, 1987). The intensity of emotional-feeling states predicted memory for the reception context in various studies (Bohannon, 1988; Bohannon & Symons, 1992; Christianson; Finkenauer et al.; Pillemer, 1984; Rubin & Kozin).

Rehearsal is also of interest as a critical predictor of later flashbulb memories (Bohannon & Symons, 1992; Christianson, 1989; Finkenauer et al., 1998;

McCloskey, Wible, & Cohen, 1988; Neisser, 1982; Wright, 1993) that includes long-term cognitive and social aspects. The cognitive aspects involve ruminations in which thoughts related to an emotional event repeatedly enter consciousness (Martin & Tesser, 1989, 1996; Tait & Silver, 1989). The social aspects consist of an urge to communicate with other people about the emotional circumstances and the experienced feelings and reactions (Luminet, Bouts, Delie, Manstead, & Rimé, 2000; Rimé, Finkenauer, Luminet, Zech, & Philippot, 1998; Rimé, Mesquita, Philippot, & Boca, 1991). When collective events are involved, rehearsal also operates through following reports in the media (television, radio, newspapers, the internet).

The rememberer's prior knowledge related to the news is also of central importance for later flashbulb memories (Conway et al., 1994; Finkenauer et al., 1998). A second, related cognitive structure is attitudes, which affect the way people approach and react to situations. The results of many studies have shown that people pay attention to information that fits with their previous expectations about a situation (e.g., Cohen, 1981), and that people's expectations also guide their later retrieval of that information (e.g., Hirt, Erickson, & McDonald, 1993). Prior knowledge and attitudes are predicted to facilitate the organization and assimilation of the incoming information into existing semantic structures in memory (Conway et al.; Finkenauer et al.).

Finally, memory for event-related facts may be encoded simultaneously with memory for the reception context (Finkenauer et al., 1998). Finkenauer et al. suggested that when people learn about an original event, all activated information—including the reception context, sensory information, and the original event—is encoded in memory (Tulving & Kroll, 1995). Thus, both the original event and the reception context will be encoded immediately. Subsequently, during rehearsal of the original event, the reactivation of that information in memory would spread to any associated information and thus to flashbulb memories. As described by Bohannon (1988), people's memory of event details may not be affected by the same factors that affect flashbulb memories for personal circumstances (see also Larsen, 1992).

We turn now to predictions about the data. Because of the large sample sizes, differences were likely to be observed for most of the variables. Thus, we used three levels of effect sizes (small, moderate, and large) to categorize the differences. As recommended by Cohen (1988), small effect sizes were when η^2 was smaller or close to .01, moderate effect sizes were when η^2 was close to .06, and large effect sizes were when η^2 was around .14.³

First, we predicted a small difference between U.S. and non-U.S. groups for memory for the reception context. Overall high levels were expected, owing to the nature of the events of 9/11 and the short delay between the events and the administration of the questionnaire. In addition, we expected very high average levels across countries (thus small differences) for the appraisal of novelty, as no similar event had ever occurred before. Researchers have shown that novelty

directly influences the emotional state of surprise, which in turn directly strengthens memory for the reception context (Finkenauer et al., 1998).

For the other variables, however, we predicted larger differences between U.S. and non-U.S. respondents. Most important, we predicted large group differences for the ratings of importance. Americans were more likely to have known someone who was killed in the attacks or to know people who were otherwise directly involved. Also, the attacks had a direct impact on daily life in America (e.g., security at public events and airports, no airline transportation for several days). If one follows the emotional-integrative model (Finkenauer et al., 1998, see the discussion for a more complete account of the model), the differences in importance–consequentiality should be reflected in variables that are supposed to be direct outcomes. Thus, we predicted moderate to large group differences for valenced emotional-feeling states, rehearsal, and memory for event-related facts. Because Americans were likely to be more knowledgeable about the WTC and the Pentagon before the attacks, we also predicted moderate differences for background knowledge. We predicted moderate differences for prior knowledge of terrorist groups as well, because Americans were more motivated to have sought out information related to the perpetrators of the attacks.

The large number of non-U.S. groups allowed us to test for differences among them in memory for reception context and for the other target variables. No a priori hypotheses were formulated for those comparisons.

In summary, we examined the immediate responses of 3,665 people from nine countries to the September 11 attacks. This study differs from previous work in its emphasis on initial reactions (with the goal of an empirical snapshot of the world's reactions) and in the large number of comparison groups, which were not limited to Western respondents.

Method

Participants

A total of 3,665 participants from nine different countries participated in the study. The participants resided in Belgium ($n = 546$), France ($n = 62$), Italy ($n = 296$), Japan ($n = 683$), Romania ($n = 716$), Switzerland ($n = 55$), The Netherlands ($n = 348$), Turkey ($n = 191$), and the United States ($n = 768$). The demographic characteristics of each national group are displayed in Table 1.⁴

The differences between nationalities were analyzed with a one-way analysis of variance (ANOVA) for age and chi-squared analyses for gender and student status. Because of the large number of tests, the Bonferroni's correction was performed on the alpha level, which was set at .005. Age significantly differed across nationalities, $F(8, 3637) = 47.2, p < .001$. Post hoc Tukey honestly significant difference comparisons revealed that the Romanian respondents were significantly older than were respondents from the other countries, with the exception of those

TABLE 1. Sample Size, Percentage of Total Sample, Mean Age, Sex, and Status by Country of Residence

Country of residence	<i>n</i>	% of total sample ¹	Age		Sex ²		Status ²	
			<i>M</i>	<i>SD</i>	Male	Female	Student	Other
Belgium	546	14.9	28.6	13.1	31.7	67.9	44.7	39.9
France	62	1.7	29.6	10.5	25.8	71.0	29.0	58.1
Italy	296	8.1	29.6	8.7	24.4	75.3	43.9	46.0
Japan	683	18.6	22.7	8.6	36.2	63.8	86.5	11.2
Romania	716	19.5	33.0	16.1	29.5	70.4	37.6	52.2
Switzerland	55	1.5	27.3	8.6	41.8	56.4	47.3	47.2
The Netherlands	348	9.5	26.4	11.0	39.7	60.3	42.0	48.8
Turkey	191	5.2	20.9	3.8	17.3	82.7	96.9	3.1
United States	768	20.9	25.9	10.2	31.9	68.1	59.9	35.9

Note. Percentages for Sex and Status do not always add up to 100 because of missing values.

¹*N* = 3,665. ²Expressed as percentage of *n*.

from France and Switzerland ($p < .001$). In addition, the Japanese and Turkish respondents were significantly younger than all other groups ($p < .001$) and did not differ from each other.

In regard to gender, the male–female ratios differed significantly across nationalities, $\chi^2(8, N = 3665) = 47.4, p < .001$. As can be seen in Table 1, the Turkish, Italian, and French groups had a relatively high percentage of women. Chi-square tests were run to compare countries two at a time. The results revealed that the gender distribution in the Turkish group significantly differed from that in the other countries, all $\chi^2(1, N > 245) > 11.3, p < .001$, with the exception of Italy and France. In addition, the Italian sample significantly differed from the Japanese and the Dutch groups, $\chi^2(1, N = 979) = 12.9, p < .001$, and $\chi^2(1, N = 544) = 16.8, p < .001$, respectively.

Furthermore, in regard to the percentage of students, a significant overall difference between nationalities emerged, $\chi^2(8, N = 3665) = 496.6, p < .001$. Chi-square tests were run to compare countries two at a time. The Turkish sample had the highest percentage of students compared with all other countries, all $\chi^2(1, N > 245) > 11.6, p < .001$. Likewise, Japan differed from all other countries, all $\chi^2(1, N > 737) > 11.6, p < .001$. Furthermore, the U.S. sample had a relatively large percentage of students and differed significantly from the other countries, all $\chi^2(1, N > 829) > 14.9, p < .001$, except from Switzerland.

Measures⁵

Memory for the reception context. Eight questions assessed recall of the circumstances in which the participants first learned about the terrorist attacks. The questions were similar to those used in previous research (Bohannon, 1988; Brown & Kulik, 1977; Conway et al., 1994; Finkenauer et al., 1998). The participants were asked the following questions about when they first heard the news of the attacks: (a) the date (day, month, year), (b) the day of the week, (c) the time of the day, (d) informant (TV, radio, newspaper, another person), (e) the place they were in (country, state, city, exact location—in the kitchen, on the street, at the office), (f) the other people present, (g) their ongoing activity, and (h) the degree of change in their ongoing activity. The questions were open-ended and judges scored the responses. For most items, responses received a 1 if the respondents provided an answer and a 0 if they did not. For items (a), (c), and (g) the scores ranged from 0 to 2.⁶ Those scores were summed together, with the maximum possible score for the variable being 11.

Novelty. On the basis of Curci et al. (2001), the participants rated five items about how (a) common, (b) unusual, (c) special, (d) strange, and (e) ordinary the event was for them. The items ranged from 1 (*not at all*) to 7 (*very much*). The novelty score was obtained by averaging the five items after scores on items (a) and (e) were inverted.

Surprise. On the basis of Curci et al. (2001), the participants rated the extent to which (a) they felt surprised, (b) the news was expected, and (c) the news was unbelievable. The items ranged from 1 (*not at all*) to 7 (*very much*). Owing to reliability problems (see Results, and sections on Measurement Issues and Structural Analyses) only item (c) was used in the analyses as the measure of surprise.

Importance–consequentiality. The participants rated four items on the extent to which the event was important (a) for themselves, (b) for their family and friends, (c) for their country, and (d) at the international level (from 1 = *not at all important*) to 7 = *very important*). In addition, they rated the extent to which the original event had consequences for them on a scale ranging from 1 (*no consequences at all*) to 7 (*many consequences*). The importance–consequentiality score was computed by averaging the four items related to importance and the single item related to consequences. All items had been used by Curci et al. (2001) and Finkenauer et al. (1998), except items (c) and (d), which were added for the present study.

Emotional-feeling state. This measure consisted of four items selected from Curci et al. (2001).⁷ The participants rated the extent to which they were upset by the news on a scale ranging from 1 (*not at all upset*) to 7 (*very much upset*). In addition, they rated the impact of the news by assessing the extent to which they felt (a) shaken, (b) affected, and (c) indifferent on a 7-point scale (from 1 = *not at all* to 7 = *very much*). The scores were averaged to form a single index of emotional-feeling state, after the score on the indifferent item was inverted.

Social sharing. Preliminary reports indicated a ceiling effect in rehearsal in the days immediately after the attacks. For instance, immediately after the attacks, almost all Americans reported discussing them more than 15 times. To avoid a ceiling effect, we modified the questions that assessed social sharing, following the mass media, and rumination; in our version, they referred to the 3 days before they completed the questionnaire.

Social sharing was assessed by six items identical to those used by Curci et al. (2001). The respondents described on 5-point scales (a) how many times they had discussed the events (from 1 = *never* to 5 = *more than 15 times*), (b) how long after hearing the news they first spoke about it (from 1 = *less than 10 minutes after* to 6 = *never*), and (c) with how many people they had discussed the news (from 1 = *none* to 5 = *more than 15*). They also used three 5-point scales (from 1 = *never* to 5 = *very often*) to rate the frequency of discussion broken down by the content of the speech (a) the event itself, (b) their feelings, and (c) the circumstances in which they heard the news. The scores were averaged to form a single index of social sharing.

Following the mass media. The participants used four 5-point scales (from 1 = *never* to 5 = *more than 15 times*) to rate how often in the previous 3 days they had followed the news via (a) TV, (b) radio, (c) newspapers, and (d) the internet.

The first three items were identical to those used by Finkenauer et al. (1998) and by Curci et al. (2001), and the fourth was added because we hypothesized that many people followed the news through that channel as well.

Rumination. The participants rated the frequency over the 3 days before they completed the questionnaire of thoughts, memories, or images related to the event on a single 5-point scale ranging from 1 (*never*) to 5 (*more than 15 times*). A similar scale was used by Curci et al. (2001).

Background knowledge. The participants answered 11 questions selected by the authors as representing general knowledge related to the September 11 attacks. We believed that it was important to assess the knowledge relating to the targeted buildings (the World Trade Center and the Pentagon) and to terrorism because those two aspects of knowledge should be associatively linked to memory for event-related facts and the reception context.⁸ The first 9 questions were related to the WTC and the Pentagon, and the last 2 questions were related to terrorism. The participants estimated (a) the number of floors in the two WTC towers, (b) the height of each of the towers, (c) the year in which the construction of the towers was completed (d) how long it took to build the towers, (e) the year in which the WTC had been previously attacked (f) when the Pentagon was built, (f) the name of the administrations located at the Pentagon, (g) the number of people who worked in the Twin Towers, and (h) the number of people who worked in the Pentagon each day. Then the respondents listed (i) terrorist attacks that had taken place in the United States since 1998, and (j) the names of terrorist groups that claim to be related to Islam.

Each answer was scored as a 2 if it was completely right, a 1 if it was basically but not entirely correct, or a 0 when no answer was given or when the answer was completely wrong. To illustrate that, consider answers to when the WTC had previously been attacked, a 2 was given to an answer of 1993, a 1 to answers of 1991, 1992, 1994, or 1995, and 0 to all other years or if no year was provided. The criteria for attributing the scores were based on a priori decisions and on an analysis of the frequency of responses. In other words, the scorings did not require subjective ratings, and thus the presence of multiple judges and the computation of interrater reliability were not involved. Scores were summed, resulting in total scores from 0 to 18 for knowledge related to the WTC and the Pentagon and from 0 to 3 for knowledge related to terrorist acts.

Attitudes. The non-U.S. respondents rated on two 7-point scales (from 1 = *not at all* to 7 = *very much*) to what extent they liked (a) the United States as a country, and (b) U.S. citizens. The two measures were specifically developed in the context of the present study. The attitude score involved averaging the two items.

Memory for event-related facts. Memory for the event involved nine questions concerning (a) the exact date of the terrorist attacks on the WTC and the Pentagon

(date, month, and year), (b) the day of the week it happened, (c) the time of the day it happened, (d) the number of planes hijacked, (e) the airline companies involved in the attacks, (f) the number of people who were aboard the planes, (g) the time at which the Twin Towers were hit, (h) the time at which the Pentagon was hit, and (i) the time lag between impact and the collapse of each of the towers. That number of items is an increase on the number of items used in other studies (Curci et al., 2001; Finkenauer et al., 1998). We were able to have more items because the target events were multifaceted (involving two different locations and four different airplanes), whereas previous events revolved around a single core fact. By having more items, we aimed to increase the content validity of the measure.⁹ Again, an answer was scored as a 2 if it was completely right, a 1 if it was basically but not entirely correct, or a 0 when no answer was given or when the answer was far from reality. Scores of 1 and 2 were defined a priori based on the distribution of the answers across subgroups. (The complete coding system is available from the first author). The scores were summed for a total score, which ranged from 0 to 20.

Procedure

The questionnaires were distributed among university students, university faculty and staff, and experimenters' acquaintances. Overall, students completed the questionnaires in groups (i.e., during class), whereas other participants were tested individually. Exceptions were the Swiss and French students, who were tested individually. All the non-U.S. participants completed a paper-and-pencil version of the questionnaire. In an effort to increase access to U.S. citizens, we developed an online version of the questionnaire using exactly the same questions as the paper-and-pencil version. Seventy-two U.S. respondents completed the paper-and-pencil version of the questionnaire, and 696 Americans completed the online version. Table 2 summarizes the periods during which the questionnaires were distributed in the different countries.

Results

Measurement Issues and Structural Analyses

We used the Statistical Package for Social Sciences Categorical Principal Component Analysis (CatPCA) procedure to aggregate each set of indicators (memory for the reception context, novelty, surprise, importance, emotional-feeling states, rehearsal, background knowledge, attitudes, memory for event-related facts) in single composite scores. CatPCA is an exploratory analysis aimed at identifying the latent dimensions underlying a set of categorical variables.¹⁰

Separate analyses were run on each set of indicators, and, for each set, the one-dimensional solution was considered. The purpose was to aggregate each set

of indicators into one summary measure so that the new measure represented the observed data with as little loss of information as was possible. Table 3 reports the eigenvalues, percentages of explained variance, and Cronbach's alphas for the considered solutions. The percentages of explained variance ranged from 30% to

TABLE 2. Period of Questionnaire Distribution by Country of Residence

Country	<i>n</i>	Period	Delay in days from event	
			<i>M</i>	<i>SD</i>
Belgium	546	Sept. 15–Oct. 3	9.98	3.17
France	62	Sept. 17–Oct. 25	10.19	5.40
Italy	296	Sept. 16–Oct. 23	8.64	5.58
Japan	683	Sept. 17–Oct. 6	17.15	5.90
Romania	716	Oct. 8–Oct. 15	30.13	1.75
Switzerland	55	Sept. 17–Sept. 21	9.13	1.24
The Netherlands	348	Sept. 14–Oct. 21	9.07	5.00
Turkey	191	Sept. 26–Oct. 30	17.81	3.50
United States				
Paper-and-pencil version	129	Sept. 17–Oct. 15	12.22	6.52
Online version	636	Oct. 1–Oct. 24	29.21	6.29

TABLE 3. CatPCA Eigenvalues, Explained Variances, and Cronbach's Alphas

Composite variables	Eigenvalues	% variance	Cronbach's alpha
Flashbulb attributes	2.43	30.31	.672
Memory for the event	3.75	37.50	.815
Background knowledge about WTC and Pentagon	4.46	44.55	.862
Background knowledge about terrorist acts	1.25	62.51	.400
Attitude	1.69	84.26	.813
Novelty	2.11	42.28	.659
Surprise	1.43	47.59	.449
Importance	2.65	53.02	.779
Negative emotional-feeling state	2.32	58.02	.759
Social sharing	2.28	38.02	.674
Following the mass media	1.85	46.22	.612

Note. WTC = World Trade Center. CatPCA = Categorical Principal Component Analysis.

84%. The lowest percentages were found for the composite variables that assessed memory for the reception context (30.31%), memory for event-related facts (37.50%), and social sharing (38.02%). The reliabilities of the scales, assessed through the Cronbach's alpha coefficients, exceeded .60, except for background knowledge about terrorist acts (.400) and surprise (.449).

The percentage of explained variance can be considered as a measure of how well the underlying construct was captured by the scales. With respect to the variables assessing memory for the reception context, memory for event-related facts, and social sharing, the low percentages of explained variance could be an effect of the dimensionality of the scales. Indeed, a two-dimensional solution for those scales will result in an increase in the amount of explained variance of between 17% and 21%. However, two-dimensional solutions would not meet the requirement of aggregating a set of observed measures into a single composite score. As a consequence, despite the low explained variance, for the purposes of the present article, only one-dimensional solutions were considered.

The variables assessing background knowledge about terrorist acts and surprise were found to have a low reliability as assessed through the Cronbach's alpha coefficients. In regard to background knowledge about terrorist acts, that may be owing to the fact that the indicator variables referring to that construct were quite heterogeneous. Indeed, one assessed knowledge of previous terrorist attacks in the United States, whereas the other required listing other terrorist groups that claim to be related to Islamic faith. In contrast, the low reliability is a more serious problem for the indicator variables that assess surprise, given the key role of surprise in flashbulb memory formation. A repeated-measures ANOVA, run on the three indicator variables assessing surprise, revealed that there was an effect of the item used, $F(2, 7090) = 395.48$, $p < .001$, $\eta^2 = .100$, thus in the subsequent analyses only the third indicator variable was used as a measure of surprise; that is, the participants' rating of how unbelievable they found the event upon first hearing of it.

CatPCA is a statistical tool used to replace the Weighted Attribute Scores (WAS) scoring system, a procedure that was first used by Neisser and Harsch in their study on flashbulb memories for the *Challenger* space shuttle disaster (Neisser & Harsch, 1992; Conway et al., 1994). They assumed that the attributes of the reception context were not all equally important. Some attributes were defined as major because they seemed to be essential to identifying the original event. Other attributes were defined as minor because one could be wrong on some details but still essentially right about what really happened. The WAS procedure involves assigning a score for each recalled major attribute plus a bonus point when the participant scores above a given threshold on the set of minor attributes. In other words, the procedure requires assigning different weights to the attributes of the reception context before they are added up in a summary measure.

The logic of CatPCA is the same as that of the WAS procedure in that a composite measure combines the scores of different indicator variables. A respondent's

score on a given indicator variable is weighted with respect to the distribution of frequencies for that indicator variable (Greenacre, 1993). Thus, CatPCA composite scores are the combinations of weighted scores on a set of indicator variables. With respect to WAS, the main advantage of CatPCA is that the weights assigned to the scores are not decided a priori by the researcher, but that they come from the empirical distributions of the response categories of the indicator variables.¹¹

CatPCA starts from a raw data matrix in which the rows are the respondents and the columns are the indicators of each considered set of variables (van de Geer, 1993; Weller & Romney, 1990). Optimal scores are derived for each row of the matrix so that they should have maximum correlations with the variables in the data matrix to be analyzed. Optimal scores are standardized coefficients, and, compared with ordinary correlation coefficients, they may be positive or negative in conformity with the direction of the correlation of the corresponding observed scores with the dimensional solution. Optimal scores do not tell anything about the average intensity of a given construct. They only indicate the position of respondents along a continuum represented by the latent construct. That continuum has the mean for the whole sample of respondents as the 0-point, thus a positive score indicates that the respondent scored above the mean for the whole sample on a given construct, a negative score indicates that the respondent scored below the mean. For example, a positive optimal score on memory for event-related facts means that the respondent (or group) answered more questions correctly about memory for the event than did a respondent (or group) who (that) had a negative score.

To give a rough picture of the mean intensity of each latent construct, the last column of Table 6 reports the average sums of scores on the indicator variables assessing memory for the reception context, memory for event-related facts, background knowledge about the WTC and the Pentagon, and background knowledge about terrorist acts for the whole sample of respondents. The same table also contains the means for each set of indicator variables assessing novelty, surprise, importance, emotional-feeling states, social sharing, and following the mass media for the whole sample of respondents. Those composite raw scores exhibit large correlation coefficients with the composite scores coming from the CatPCA procedure (from .71 to .96) and thus can be considered as informative of the mean for the whole sample on each latent construct.

The saved optimal scores resulting from the CatPCA analyses were used as final measures for each set of constructs and were analyzed in one-way ANOVAs by citizenship to test whether memory for the reception context, memory for event-related facts, background knowledge, attitudes, emotional reactions, and rehearsal differed across U.S. and non-U.S. citizens. For rumination and surprise, we directly entered the raw scores in the analysis, because only one questionnaire item assessed reliably those dimensions.¹²

Because of the high number of statistical tests, for both main effects and planned contrasts and to reduce the possibility of Type I errors, we adopted a more

conservative approach by using the Bonferroni's correction to lower the p level to .005. As a result, comparisons significant at $p > .005$ level were not considered.

Descriptive Analyses

Before describing the results of the CatPCA analyses, we will provide the reader with some descriptives to aid the understanding of the large data set. Table 4 shows the percentages of U.S. and non-U.S. respondents who answered the questions about memory for reception context. Those percentages are very high for both groups of respondents, although slightly higher for the U.S. respondents. They range from 85.2% for the variable assessing other people present to 99.1% for that assessing location.

Table 4 also shows the percentages of respondents who answered correctly the questions about event-related facts. Those percentages range from 44.6% to 91.7%. The participants were least able to report the exact time at which the Pentagon was

TABLE 4. Percentages of Correct Answers for Memory for the Reception Context and Correct Answers for Memory for Event-Related Facts

Variable	Respondents		Total sample
	U.S.	Non-U.S.	
<i>Memory for reception context</i>			
Date	96.5	93.5	94.2
Day of week	97.4	94.6	95.2
Time of day	97.9	96.3	96.6
Informant	98.8	98.5	98.6
Location	99.5	99.0	99.1
Other people present	86.7	84.8	85.2
Ongoing activity	99.2	98.0	98.3
Changes in activity	97.2	92.1	93.2
<i>Memory for event-related facts</i>			
Date	97.2	90.3	91.7
Day of week	97.1	89.1	90.7
Time of day	79.8	72.2	73.8
Number of planes	97.6	86.5	88.9
Airline companies	88.6	48.8	57.2
Number of people on planes	67.7	49.1	53.0
Time Twin Towers were hit	91.1	76.7	79.7
Time Pentagon was hit	47.0	44.0	44.6
Time collapse South Tower	70.3	50.8	54.9
Time collapse North Tower	64.1	52.3	54.8

hit, the exact time of the collapse of the WTC towers, the names of the airline companies, and the number of people on the hijacked planes. The percentages of responses were again higher for the U.S. respondents.

Table 5 reports the percentages of correct responses for questions probing background knowledge. For background knowledge about the WTC and the Pentagon, correct answers varied from 13.2% (how many years it took to build the WTC) to 59% (the number of floors in the WTC). The U.S. respondents scored higher than did the non-U.S. respondents on all variables, except those assessing the height of the towers and the number of people working at the WTC and the Pentagon. Overall, the respondents were less able to answer correctly the questions about terrorist groups (31.6%) and terrorist acts (15%), but the U.S. respondents exhibited higher rates of correct answers on both indicator variables assessing background knowledge about terrorist acts.

Structural Analyses by Citizenship

Differences within the U.S. sample. First, the sample of U.S. respondents was split into three groups (West coast, central regions, East coast), to test whether location within the United States affected the results. One-way ANOVAs were run on the CatPCA optimal scores for the U.S. respondents, with location (West coast

TABLE 5. Percentages of Correct Answers for Background Knowledge About the World Trade Center, the Pentagon, and Terrorist Acts

Background knowledge	Respondents		Total sample (<i>N</i>)
	U.S.	Non-U.S.	
About WTC			
Number of floors	72.0	55.5	59.0
Height of Tower 1	24.1	34.6	32.4
Height of Tower 2	23.4	31.5	29.9
No. people working there	21.3	38.9	35.1
Year completed	24.6	15.8	17.6
Time taken to build	15.7	12.5	13.2
Previous attack	59.2	31.9	37.6
About Pentagon			
Year completed	24.7	10.4	13.3
Administration	65.2	56.8	58.6
No. people working there	12.4	14.4	13.9
About acts of terrorism			
Previous attacks on U.S.	25.5	12.2	15.0
Terrorist groups	31.8	31.6	31.6

Note. WTC = World Trade Center. *N* = 3,665.

vs. central regions vs. East coast) as the between-participants factor. Significant differences were only found for importance, $F(2, 529) = 5.26, p = .005, \eta^2 = .020$, background knowledge about the WTC and the Pentagon, $F(2, 529) = 4.20, p = .005, \eta^2 = .020$, and about terrorist acts, $F(2, 529) = 6.77, p < .001, \eta^2 = .025$. The respondents from the West coast reported higher ratings of importance than did those from the central regions of the United States ($p < .005$). The respondents from the East coast scored higher on background knowledge about the WTC and the Pentagon and about terrorist acts than did those from the West coast. The respondents from the East coast also scored higher on the last variable than did those from the central regions of the United States ($ps < .005$). Thus, the U.S. sample was very homogeneous for most of the variables investigated and was considered as a whole (see Pezdek, 2003, for contrasting results on this issue).

Differences between U.S. and non-U.S. citizens. To compare responses from the U.S. and the non-U.S. participants, one-way ANOVAs were run on the CatPCA optimal scores for each set of indicators, with citizenship (U.S. vs. non-U.S. respondents) as the between-participants factor, and gender and elapsed time between September 11, 2001 and questionnaire completion as covariates (see Table 6). All comparisons were significant at .001 alpha level. Thus, the U.S. respondents scored higher than did the non-U.S. respondents on all considered variables. Large effect sizes (η^2 around .14) were found for ratings of importance and memory for event-related facts. Moderate effect sizes (η^2 around .06) were found for background knowledge (about the WTC, the Pentagon, and terrorist acts), and, to a lesser extent, for emotional-feeling states. Finally, small effect sizes (η^2 around .01) were observed for all the other variables (memory for the reception context, novelty, surprise, social sharing, following the mass media, and rumination).

Differences among non-U.S. citizens. The final step was to consider differences among the non-U.S. respondents. One-way ANOVAs were run on the CatPCA optimal scores, with citizenship (French-speaking countries, Italy, Japan, Romania, The Netherlands, Turkey) as the between-participants factor, and covering the gender and time elapsed between September 11, 2001 and the completion of the questionnaire. Table 7 displays the results of those comparisons. All comparisons were significant at .001 alpha level, except that memory for the reception context was significant at .005 alpha level.

The largest effect size was for emotional-feeling states ($\eta^2 = .179$). That effect was explained by high scores in French-speaking countries, Italy, The Netherlands, and Romania, whereas scores for Japan and Turkey were very low. The other large effect size was for surprise ($\eta^2 = .137$). Again, that was driven by lower levels of surprise in the Turkish respondents.

For the two variables assessing background knowledge, effect sizes were in between moderate and large ($\eta^2 = .095$ and .103, respectively). It is noteworthy that the Turkish respondents showed different patterns for those two variables.

TABLE 6. ANOVAs by Citizenship (U.S. vs. Non-U.S.) on the CatPCA Optimal Scores

Composite variables	<i>F</i>	<i>df</i>	η^2	Citizenship				Raw scores of total sample		Range
				U.S.		Non-U.S.		<i>M</i>	<i>SD</i>	
				<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Memory for reception context ¹	12.72*	1, 3268	.004	0.080	1.120	-0.017	0.988	10.460	1.030	0-11
Novelty	36.17*	1, 3432	.010	0.191	0.922	-0.054	1.014	5.910	1.030	1-7
Surprise ^{1,2,3}	54.41*	1, 3252	.016	5.890	1.750	5.460	1.880	5.550	1.860	1-7
Importance ²	539.00*	1, 3597	.130	0.688	0.812	-1.185	0.965	5.480	1.030	1-7
Emotional-feeling state ^{1,2}	152.02*	1, 3368	.043	0.375	0.982	-0.143	0.968	5.370	1.290	1-7
Social sharing ¹	110.63*	1, 3203	.037	0.114	0.999	-0.047	0.994	3.480	0.680	1-6
Following the mass media ^{1,2}	49.71*	1, 3178	.013	0.115	1.116	-0.038	0.962	2.100	0.730	1-5
Rumination ^{1,3}	61.54*	1, 3221	.019	2.980	1.230	2.900	1.170	2.930	1.190	1-5
Background knowledge										
WTC, Pentagon ^{1,2}	216.79*	1, 3263	.062	0.323	0.924	-0.087	1.012	4.620	3.680	0-18
Acts of terrorism ^{1,2}	249.37*	1, 3263	.070	0.444	1.070	-0.119	0.943	0.600	0.840	0-3
Memory for event-related facts ^{1,2}	523.27*	1, 3263	.138	0.498	0.601	-0.165	1.056	11.73	4.020	0-20

Note. CatPCA = Categorical Principal Component Analysis.

¹Results were controlled for delay from the event, significant at $p < .001$. ²Results were controlled for gender, significant at $p < .001$. ³Raw scores, with range 1-7.

* $p < .001$.

That is, they scored below average for background knowledge related to the WTC, but highest for knowledge about terrorism.

Several moderate-size effects were observed; many of the differences were driven by lower levels in the Turkish respondents. A moderate effect size ($\eta^2 = .073$) was obtained for attitudes toward the United States, with lowest scores in the Turkish and in the Dutch groups. The effect size was moderate ($\eta^2 = .047$) for memory for event-related facts, with planned contrasts revealing the lowest levels among the Romanian and Japanese respondents. For the three rehearsal variables (social sharing, rumination, following the mass media), the effect sizes were in the small to moderate ranges (η^2 from .022 to .047, respectively). It is interesting to note that Romania and Turkey had the lowest levels on these three rehearsal variables. For novelty effect size was in the small to moderate range ($\eta^2 = .036$), and again the lowest level was observed among the Turkish respondents.

Effect sizes were small for importance ($\eta^2 = .018$) and memory for reception context ($\eta^2 < .01$), meaning that those variables varied little as a function of citizenship.

Discussion

Attacks such as those of September 11, 2001 had never happened before, and consequently people around the world rated the attacks as *novel* and *surprising*, and rehearsed the events at high levels. Across the globe, people remembered how they first heard the news. To borrow a term used typically to describe the beginning of the American Revolution in 1776, the events of 9/11 were “a shot heard round the world.”

Only very small differences emerged across countries in memory for reception context (the effect size was below 1%). On average, more than 95% of the U.S. respondents answered the questions about how they heard the news, and similarly more than 90% of the non-U.S. sample did so. That is, regardless of citizenship, the respondents reported remembering most of the details of how they heard the news. That high level of immediate memory is similar to that which has been observed in other studies (e.g., Christianson & Engelberg, 1999; Conway et al., 1994; Curci et al., 2001; Pezdek, 2003; Schmolck et al., 2000; Wright, 1993).

A major strength of our study was the number of comparison groups we had. That is, in addition to assessing the United States and Western European countries (as have the other published studies on group differences in flashbulb memories), we also assessed citizens of Japan, Romania, and Turkey. That was crucial. As the observant reader would have noticed, many of the differences within non-U.S. countries were driven by the Turkish respondents. They reported particularly low levels of novelty, surprise, emotional-feeling states, social sharing, rumination, following the media, attitudes, and background knowledge about the WTC and the Pentagon.

Overall, that pattern of results makes sense given Turkish history and culture. That is, Turkey has fallen victim to a number of terrorists attacks in the last few

TABLE 7. ANOVAs by Citizenship (Non-U.S. Countries) on the CatPCA Optimal Scores

Composite variables	<i>F</i>	<i>df</i>	η^2	Citizenship					
				French-speaking countries	Italy	Japan	Romania	The Netherlands	Turkey
Memory for reception context	3.394*	5, 2891	.006						
<i>M</i>				.006	.002	-.041	-.119 _a	.108 _a	.090
<i>SD</i>				.980	1.262	.944	.924	.987	.512
Novelty	19.86**	5, 2663	.036						
<i>M</i>				-.089 _a	.044 _b	.127 _{ac}	.002 _d	-.212 _{ce}	-.644 _{abcde}
<i>SD</i>				.960	.843	1.102	.856	1.159	1.008
Surprise ^{1,3,4}	79.91**	5, 2520	.137						
<i>M</i>				5.987 _a	5.865 _b	5.073 _{abc}	5.607 _{acd}	5.773 _{ce}	3.268 _{abcde}
<i>SD</i>				1.544	1.743	1.871	1.886	1.511	1.985
Importance ³	10.33**	5, 2827	.018						
<i>M</i>				-.293 _a	.141 _{ab}	.138 _{ab}	-.148 _b	-.309 _b	-.232 _b
<i>SD</i>				.906	.890	.973	1.075	.822	.920
Emotional-feeling state ³	113.5**	5,2603	.179						
<i>M</i>				.128 _a	.506 _{ab}	-.694 _{abc}	.172 _{bcd}	-.110 _{bcde}	-.653 _{abde}
<i>SD</i>				.956	.875	.677	1.024	.897	.693
Social sharing ²	23.40**	5,2473	.047						
<i>M</i>				.230 _a	.488 _{ab}	-.254 _{abc}	-.367 _{abd}	.371 _{cde}	-.560 _{abce}
<i>SD</i>				.828	.880	.957	1.085	.799	.890

Following the mass media ^{2,3}	11.01**	5,2446	.022						
<i>M</i>				-.128 _a	.225 _{ab}	.066 _{ac}	-.285 _{bcd}	.276 _{acde}	-.250 _{bce}
<i>SD</i>				.915	1.045	.997	.870	1.056	.645
Rumination ^{2,4}	24.56**	5,2520	.046						
<i>M</i>				2.930 _a	3.402 _{ab}	2.766 _{bc}	2.606 _{abd}	3.587 _{acde}	2.410 _{abce}
<i>SD</i>				1.192	1.040	1.081	1.133	1.195	.935
Background knowledge about WTC, Pentagon ^{2,3}	57.86**	5,2530	.103						
<i>M</i>				.530 _a	-.410 _{ab}	-.286 _{ac}	-.475 _{acd}	.293 _{abcde}	-.210 _{ae}
<i>SD</i>				.884	.808	1.092	.842	.870	.917
Background knowledge of terrorist attacks ^{2,3}	53.27**	5,2530	.095						
<i>M</i>				.264 _a	-.471 _{ab}	-.335 _{ac}	-.351 _{ad}	.115 _{bcd}	.427 _{abcd}
<i>SD</i>				1.088	.792	.831	.823	1.020	.503
Attitude ^{3,5}	44.60**	5,2833	.073						
<i>M</i>				-.175 _a	.104 _{ab}	.132 _{ac}	.310 _{ac}	-.492 _{abc}	-.326 _{bc}
<i>SD</i>				.793	1.006	.954	1.212	.686	.824
Memory for event-related facts ^{2,3}	24.76**	5,2530	.047						
<i>M</i>				.384 _a	-.088 _{ab}	-.385 _{abc}	-.630 _{abcd}	.365 _{bcde}	-.075 _{acde}
<i>SD</i>				.720	.980	1.230	1.024	.693	.688

Note. Means in a row sharing subscripts are significantly different at least at .005 alpha level. For each composite variable, higher positive means indicate higher scores on that variable. French-speaking countries = France, Switzerland, and Belgium. WTC = World Trade Center. CatPCA = Categorical Principal Component Analysis.

¹Results were controlled for delay from the event, significant at $p < .005$. ²Results were controlled for delay from the event, significant at $p < .001$.

³Results were controlled for gender, significant at $p < .001$. ⁴Raw scores, with range 1–7. ⁵ M for the composite raw score for the total sample = 4.08, $SD = 1.36$ (range: 1–7).

* $p < .005$. ** $p < .001$.

years, as well as having experienced a serious earthquake. In both cases, thousands of people were killed, which likely habituated Turkish citizens to such catastrophes (and thus reducing the intensity of their responses to the 9/11 attacks). Turkey's history of terrorist attacks (and its geographical proximity to many terrorist bases) likely played a role in the Turkish participants' high level of knowledge about terrorist groups ($M = 83.8\%$ correct, compared with a sample mean of 31.6%). Cultural and economic differences help us understand the other differences. For example, the collectivistic nature of the country could explain the low scores on emotional-feeling states, social sharing, and rumination (e.g., Basabe et al., 2000, 2002; Fernandez, Carrera, Sanchez, Paez, & Candia, 2000). It has been shown that the dimension individualism-collectivism is important for predicting differences in the way people report emotional reactions (e.g., Fernandez et al.). In individualistic countries, the public reporting of one's emotional state is more socially desirable than it is in collectivistic countries, which results in average higher levels of reported emotional intensity in individualistic countries (e.g., Fernandez et al.). It has also been shown that people's focus of attention is more self-centered in individualistic countries, which may result in more rumination and social sharing (e.g., Basabe et al., 2000, 2002).

In addition, Turkey has a different religion and different values compared with Western countries, which might also have exacerbated differences. Turkey is also less economically developed, likely reducing media coverage of the events. The country has also suffered economically from United States-imposed prohibitions on trade with countries accused of helping terrorist acts immediately after the 9/11 attacks; that likely played into the less positive attitudes toward the United States held by the Turkish respondents.

On the one hand, it is interesting to note that the Turkish respondents still displayed high levels of memory for the reception context. That is, they were almost as likely as the U.S. respondents were to remember the details of having learned the news of 9/11. It may be that their knowledge of terrorism helped them to encode and assimilate news about the attacks (see Conway et al., 1994; Finke- nauer et al., 1998). On the other hand, one might have expected that all the differences between the Turkish respondents and the other respondents would have led to differences in memory for the reception context. That is, the Turkish respondents reported less novelty, less surprise, less importance, less emotional response, less social sharing and media watching—and yet they still reported very high levels of memory for the reception context. That suggests that the variables that are critical for later flashbulb memory formation are not prerequisites for immediate memory for reception context, otherwise we would not have expected to observe such high levels of memory for reception context among the Turkish respondents. Certainly, we would predict long-term differences across groups in flashbulb memories. If we had relied solely on existing cross-cultural data (e.g., Conway et al.), then we might have predicted differences in immediate memory.

Instead, the results mean that cross-cultural differences are likely to emerge

over time, which is consistent with the idea that distortions in flashbulb memory emerge over time (Schmolck et al., 2000). It is not that citizens in different countries failed to encode where they were, who told them about events, or other details of the reception context. Even though we have been labeling memory for reception context as immediate, that is not to be confused with short-term memory. That is, the events were encoded in long-term memory in the days and weeks following September 11. Those data eliminate encoding differences as a possible explanation for later predicted differences in flashbulb memories (i.e., consistency in memory for reception context over time).

In addition to telling us something about immediate reactions to and memory for September 11, the data set also yields predictions about what the respondents will remember and believe in the future. We set those predictions within the emotional-integrative model of flashbulb memory formation of Finkenauer et al. (1998). In that model, the original event (the news) is first evaluated in terms of novelty, and the outcome of that appraisal check elicits surprise, which contributes directly to the formation of flashbulb memory. That is the first and most direct path. A second, more indirect path was also hypothesized by Finkenauer et al. At the same time that novelty is appraised, importance and consequentiality are also appraised, which results in emotional-feeling states. In addition, the magnitude of the resulting emotional arousal is determined by the effects of novelty, that is, surprise. Emotional arousal subsequently affects flashbulb memory indirectly through its social aftermath. That is, greater emotionality leads to more subsequent rehearsal, as people engage in social sharing of emotions (e.g., Luminet, Bouts, et al., 2000), rumination (e.g., Luminet, Zech, Rimé, & Wagner, 2000), and follow media reports. It is important to note that people rehearse both how they heard the news and the events themselves. That rehearsing improves recall of the details of the original event, which in turn strengthens the memory for associated matters, thereby assisting in the encoding, maintenance, and retrieval of flashbulb memory. Finally, background knowledge and attitudes toward the issue affect the importance attributed to the event, the emotional reaction, and subsequent overt rehearsal.

In the present study, we found that the U.S. respondents rated the events as more important and consequential, and that they had better memory for event-related facts. To some extent, the U.S. respondents also had higher emotional-feeling states and more background knowledge. Taken together with reference to the emotional-integrative model, the results led us to predict higher levels of flashbulb memories in the U.S. respondents—but as a consequence of the indirect path. We expected the model would predict no differences in flashbulb memories on the direct path because only small differences were observed for novelty and surprise.

Interesting predictions can also be made for differences in flashbulb memories among the non-U.S. respondents. Only the Turkish group would have fewer flashbulb memories from the direct path because those respondents showed lower levels of novelty and surprise. If differences for other national groups were found, they would likely come from the indirect path. From that perspective, the results found

from Romania and Japan are very interesting. The Romanians were particularly low on memory for event-related facts, a direct predictor of flashbulb memory, and they were also low on background knowledge. The Japanese respondents reported fewer emotional reactions than did the other groups, as well as less background knowledge. On the basis of those results, Japanese and Romanian citizens are expected to have less consistent flashbulb memories, compared with Western European countries. But it is interesting to note that whereas the levels of background knowledge were quite close between those two groups, they differed substantially on emotional-feeling states (lower level in Japan) and to a lower extent on memory for event-related facts (lower level in Romania). Thus, if less consistent memory for the reception context is found at follow-up in the Japanese group than in the Romanian group, that would suggest that emotional-feeling states play a more crucial role in the formation of flashbulb memory than memory for event-related facts. The same would apply for the role of memory for event-related facts if lower levels of flashbulb memory were found for the Romanian group. Those predictions will be investigated in future follow-up studies.

Finally, two other strengths of the present study are noteworthy. First, in previous studies on flashbulb memories, researchers assumed the presence of certain variables, but did not assess them. The present study is clearly an improvement on that in the sense that we empirically assessed all the variables that have been found in the literature to affect directly or indirectly the formation of flashbulb memories. A second strength is that in many previous studies there were very low levels on some of the key variables such as personal importance (e.g., Schmolck et al., 2000; Wright, 1993) and intensity of emotion or frequency of rehearsal (e.g., Conway et al., 1994; Pillemer, 1984). The events of September 11, 2001 were tragically unique in that we would expect very high levels of variables such as importance and emotional intensity among Americans. Many Americans knew people involved in the attacks, or at least knew people who lived in the targeted cities. In addition, across the United States, people were without airline transportation for days and faced new security systems in public places. There was also a large financial toll from the attacks and the military response in Afghanistan, both of which had an impact on the economy. Whereas personal consequences were likely to reach average high levels in the U.S. group, variations were still expected because all Americans did not experience those personal consequences in the same way, depending on their professional activities or their location.

In conclusion, the results of the present study revealed that shortly after the 9/11 attacks, some substantial differences were already apparent on variables that are supposed to directly or indirectly have an impact on the formation of flashbulb memories. The U.S. respondents assessed the event as far more important, and they knew more event-related facts. The inclusion of comparison countries that differed economically and/or culturally from the United States also yielded interesting results. Within the non-U.S. respondents, major differences were observed for emotional-feeling states. Because September 11 yielded such strong responses in

the U.S. respondents, both the present data and future follow-up studies provide unique opportunities to test models of flashbulb memory formation (Conway, 2002). Because the pattern of responses varied across countries, the present data set will also offer new ways to determine which predictors are the most relevant for the formation of flashbulb memories (e.g., Brown & Kulik, 1977; Conway et al., 1994; Finkenauer et al., 1998).

REFERENCES

- Basabe, N., Paez, D., Valencia, J., Gonzalez, J. L., Rimé, B., & Diener, E. (2002). Cultural dimensions, socioeconomic development, climate, and emotional hedonic level. *Cognition and Emotion, 16*, 103–125.
- Basabe, N., Paez, D., Valencia, J., Rimé, B., Pennebaker, J. W., Diener, E., et al. (2000). Sociocultural factors predicting subjective experience of emotion: A collective level analysis. *Psicothema, 12*, 55–69.
- Bohannon, J. N., III. (1988). Flashbulb memories for the space shuttle disaster: A tale of two theories. *Cognition, 29*, 179–196.
- Bohannon, J. N., III, & Symons, V. L. (1992). Flashbulb memories: Confidence, consistency, and quality. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of flashbulb memories* (pp. 65–91). New York: Cambridge University Press.
- Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5*, 73–99.
- Christianson, S-Å. (1989). Flashbulb memories: Special, but not so special. *Memory & Cognition, 17*, 435–443.
- Christianson, S-Å., & Engelberg, E. (1999). Memory and emotional consistency: The *MS Estonia* ferry disaster. *Memory, 7*, 471–482.
- Cohen, C. (1981). Person categories and social perception: Testing some boundaries of the processing effects of prior knowledge. *Journal of Personality and Social Psychology, 40*, 441–452.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Conway, M. A. (1995). *Flashbulb memories*. Hillsdale, NJ: Erlbaum.
- Conway, M. A. (2002). *Testing the flashbulb memory hypothesis*. Unpublished manuscript, University of Bristol, England.
- Conway, M. A., Anderson, S. J., Larsen, S. F., Donnelly, C. M., McDaniel, M. A., McClelland, A., et al. (1994). The formation of flashbulb memories. *Memory & Cognition, 22*, 326–343.
- Curci, A., Luminet, O., Finkenauer, C., & Gisle, L. (2001). Flashbulb memories in social groups: A comparative test–retest study of the memory of French President Mitterrand’s death in a French and a Belgian group. *Memory, 9*, 81–101.
- Fernandez, I., Carrera, P., Sanchez, F., Paez, D., & Candia, L. (2000). Differences between cultures in emotional verbal and nonverbal reactions. *Psicothema, 12*, 83–92.
- Finkenauer, C., Luminet, O., Gisle, L., El-Ahmadi, A., Van der Linden, M., & Philippot, P. (1998). Flashbulb memories and the underlying mechanisms of their formation: Toward an emotional-integrative model. *Memory & Cognition, 26*, 516–531.
- Frijda, N. H. (1986). *The emotions*. Cambridge: Cambridge University Press.
- Gifi, A. (1990). *Nonlinear multivariate analysis*. New York: Wiley.
- Greenacre, M. J. (1993). *Correspondence analysis in practice*. London: Academic Press.
- Hirt, E. R., Erickson, G. A., & McDonald, H. E. (1993). Role of expectancy timing and outcome consistency in expectancy-guided retrieval. *Journal of Personality and Social Psychology, 65*, 640–656.

- Hornstein, S. L., Brown, A. S., & Mulligan, N. W. (2003). Long-term flashbulb memory for learning of Princess Diana's death. *Memory, 11*, 293–306.
- Larsen, S. F. (1992). Potential flashbulbs: Memories of ordinary news as the baseline. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of flashbulb memories* (pp. 32–64). New York: Cambridge University Press.
- Lazarus, R. S. (1982). Thoughts on the relations between emotion and cognition. *American Psychologist, 37*, 1019–1024.
- Leventhal, H. (1984). A perceptual-motor theory of emotion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17, pp. 117–182). New York: Academic Press.
- Luminet, O., Bouts, P., Delie, F., Manstead, A. S. R., & Rimé, B. (2000). Social sharing of emotion following exposure to a negatively valenced situation. *Cognition and Emotion, 14*, 661–688.
- Luminet, O., Zech, E., Rimé, B., & Wagner, H. L. (2000). Predicting cognitive and social consequences of emotional episodes: The contribution of emotional intensity, the Five Factor Model and alexithymia. *Journal of Research in Personality, 34*, 471–497.
- Martin, L. L., & Tesser, A. (1989). Toward a motivational and structural theory of ruminative thought. In J. S. Ulman & J. A. Bargh (Eds.), *Unintended thoughts* (pp. 306–326). New York: Guilford.
- Martin, L. L., & Tesser, A. (1996). Some ruminative thoughts. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition: Vol. 9. Ruminative thoughts* (pp. 1–48). Mahwah, NJ: Erlbaum.
- McCloskey, M., Wible, C. G., & Cohen, N. J. (1988). Is there a special flashbulb-memory mechanism? *Journal of Experimental Psychology: General, 117*, 171–181.
- Neisser, U. (1982). Snapshots or benchmarks? In U. Neisser (Ed.), *Memory observed* (pp. 43–48). San Francisco: Freeman.
- Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about Challenger. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of flashbulb memories* (pp. 9–31). New York: Cambridge University Press.
- Neisser, U., Winograd, E., Bergman, E. T., Schreiber, C. A., Palmer, S. E., & Weldon, M. S. (1996). Remembering the earthquake: Direct experience vs. hearing the news. *Memory, 4*, 337–357.
- Oatley, K., & Johnson-Laird, P. N. (1987). Toward a cognitive theory of emotion. *Cognition and Emotion, 1*, 29–50.
- Pezdek, K. (2003). Event memory and autobiographical memory for the events of September 11, 2001. *Applied Cognitive Psychology, 17*, 1033–1045.
- Pillemer, D. B. (1984). Flashbulb memories of the assassination attempt on President Reagan. *Cognition, 16*, 63–80.
- Rimé, B., Finkenauer, C., Luminet, O., Zech, E., & Philippot, P. (1998). Social sharing of emotion: New evidence and new questions. *European Review of Social Psychology, 9*, 145–189.
- Rimé, B., Mesquita, B., Philippot, P., & Boca, S. (1991). Beyond the emotional event: Six studies on the social sharing of emotion. *Cognition and Emotion, 5*, 435–465.
- Rubin, D. C., & Kozin, M. (1984). Vivid memories. *Cognition, 16*, 81–95.
- Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 293–318). Hillsdale, NJ: Erlbaum.
- Schmolck, H., Buffalo, E. A., & Squire, L. R. (2000). Memory distortions develop over time: Recollections of the O. J. Simpson trial verdict after 15 and 32 months. *Psychological Science, 11*, 39–45.
- Tait, R., & Silver, R. C. (1989). Coming to terms with major negative life events. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thoughts* (pp. 351–382). New York: Guilford.

- Tulving, E., & Kroll, N. (1995). Novelty assessment in the brain and long-term memory encoding. *Psychonomic Bulletin and Review*, 2, 387–390.
- van de Geer, J. P. (1993). *Multivariate analysis of categorical data: Applications*. Newbury Park, CA: Sage.
- Weller, S. C., & Romney, A. K. (1990). *Metric scaling: Correspondence analysis*. London: Sage.
- Wright, D. B. (1993). Recall of the Hillsborough disaster over time: Systematic biases of flashbulb memories. *Applied Cognitive Psychology*, 7, 129–138.
- Young, F. W. (1981). Quantitative analysis of qualitative data. *Psychometrika*, 46, 357–388.

NOTES

1. The percentages we report refer to the average proportion of people who were able to provide answers on questions about their contextual memory shortly after the events of 9/11 under investigation. However, such immediate answers do not indicate how accurate people were in their memory for the reception context.

2. The authors believe that no terrorist organization can be associated with religious faiths. The common message of all religions is an emphasis on psychological and social peace; no religion encourages terrorism. However, a terrorist group can gain support by alleging affiliation to an ideological representation and then use that representation to progress in nonpeaceful aims. Al Qaeda does not represent a true religious faith. Therefore, the authors avoided using the religious adjective, which is erroneously associated with the terrorist organization, Al Qaeda.

3. As noted by Cohen (1988), it is important to remember that effect sizes are conventional and arbitrary. There is thus no gold standard that can be applied to all studies. “The terms small, medium, and large are relative, not only to each other, but to the area of behavioral science or even more particularly to the specific content and research method being employed in any given investigation. In the face of this relativity, there is a certain risk inherent in offering conventional operational definitions for these terms in power analysis in as diverse a field of inquiry as behavioral science. The risk is nevertheless accepted in the belief that more is to be gained than lost by supplying a common conventional frame of reference, which is recommended for use only when no better basis for estimating the effect size is available” (Cohen, p. 25). In the field of research on flashbulb memories we are clearly in a situation in which not enough studies have been published on effect sizes, and thus it is necessary to refer to conventional frames of reference, although it is necessary to remain very careful. In terms of conventions, Cohen (see pp. 274–288, but also pp. 25–27), small effects are equivalent to $\eta^2 = .0099$, thus accounting for approximately 1% of the variance. Medium effects correspond to $\eta^2 = .0588$, thus approximately 6% of the variance, and large effects correspond to $\eta^2 = .1379$, or approximately 14% of the variance.

4. Because of low response frequencies, the respondents from Switzerland and France were aggregated with the respondents from Belgium. That resulted in a new category for the variable citizenship, corresponding to the respondents from French-speaking countries. The three countries did not seem to differ significantly on the variables considered, except for the prior knowledge about the event. The Swiss respondents had a better background knowledge about the event than did the Belgians and the French, $F(2, 660) = 7.26, p = .001$. The respondents from those three countries are thus considered together in the article and are referred to as being from “French-speaking countries.”

5. Internal consistencies (Cronbach’s alphas) are reported in Table 3.

6. For the date question, the answer was scored as a 2 when it included the year, the month, and the day the respondent heard about the news. It was scored as a 1 when it

included only the year and the month. For the hour question, the answer was scored as a 2 when an hour was mentioned, and as a 1 when only the part of the day was mentioned. For the change in ongoing activity, answers received a score of 2 when changes were mentioned in comparison with the question about the ongoing activity, and 1 when no changes were mentioned.

7. In the original questionnaire developed by Curci et al. (2001), 12 items related to emotional-feeling states were considered. To make the questionnaire shorter, only the 4 items that evidenced the highest reliability were selected for the present study.

8. Questions related to background knowledge have to be specific to the event under investigation. It is thus difficult to build them based on theoretical assumptions. For previous studies related to the death of political leaders, the questions were related to the important steps of their career (e.g., year of election as president, names of prime ministers who have worked with him, see Curci et al., 2001).

9. Although some questions on the memory for event-related facts and on the memory for the reception context were related to a similar content (e.g., time), it is important to remember that memory for event-related facts refers to the actual time the events occurred, and that the memory for the reception context refers to the time each respondent heard about it. It is possible, however, that some respondents inferred their answers to one type of memory from the answers provided for the other type.

10. The CatPCA procedure was used as an optimal scaling procedure (Young, 1981) aimed at looking for new scale values so that they are as closely correlated as possible to a set of observed scores (Gifi, 1990; van de Geer, 1993). It can be defined as a kind of principal component analysis for categorical or ordered categorical variables.

11. A WAS-like procedure was applied on the indicator variables assessing the memory for the reception context. The original procedure (Neisser & Harsch, 1992) assigned scores ranging from 0 to 2 to all indicators. Scores for place, informant, and ongoing activity were summed up. Furthermore, the respondents received an additional score of 1 if they scored at least 3 on the indicators assessing other people present and time. The final measure for flashbulb memory ranged from 0 to 7. The WAS procedure was applied on the measures of consistency. In the present study, no measures of consistency were used. However, in conformity with the WAS procedure, scores for place, informant, and ongoing activity were summed up. Furthermore, the respondents received an additional score if they scored at least 2 on the indicators assessing other people present and time. The final measure for memory for the reception context ranged from 0 to 5. The correlation between the WAS-like measure for memory for the reception context and the CatPCA composite score was found to be .751. That indicates that a correspondence exists between the two procedures, but the CatPCA scoring is to be preferred for the aforementioned specified reasons.

12. The matrix of component loadings is not reported in the manuscript but is available on request from the second author.

Manuscript received April 28, 2003

Revision accepted for publication January 14, 2004

Copyright of Journal of General Psychology is the property of Heldref Publications and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.