

Running head: FLASHBULB MEMORIES

Flashbulb Memories? The Effects of When the Initial
Memory Report was Obtained

Robert G. Winingham

Baylor University

Ira E. Hyman Jr.

Dale L. Dinnel

Western Washington University

Abstract

Why have some researchers found reports of flashbulb memories to be stable, while others have observed inconsistencies? Paradoxically, it appears that relatively long delays between event and initial documentation have produced greater consistency of participants' reports. To investigate this directly, we collected the initial documentation of hearing about O.J. Simpson's acquittal either 5 hours or one week after the acquittal was read. Observed consistency of memories varied as a function of documentation time; following an 8-week retention, the delayed reports were more consistent. The delayed group also reported fewer propositions in their initial documentation. We proposed a consolidation model to explain these results: During the days immediately following a newsworthy event, the narrative structure of these memories changes in that some details are forgotten. After this consolidation period, the memories may solidify. Thus, it may have been easier for the delayed group to provide consistent memories at the two intervals.

Flashbulb Memories?

The Effects of When the Initial Memory Report was Obtained

When asking people where they were when they heard that President Kennedy had been shot, Brown and Kulik (1977) observed long-lasting, detailed memories. They proposed that such “flashbulb memories” may be caused by a special memory mechanism that operates during the encoding of surprising, important, and emotionally arousing events. This mechanism causes people to store an exact record of the contents of the mind. The resulting memory is long-lasting, and also accurate. Most research since Brown and Kulik’s initial report has centered around their claim of extreme accuracy and resistance to forgetting (e.g. McCloskey, Wible, & Cohen, 1988; Weaver, 1993). Thus the debate concerning a special mechanism for flashbulb memories is partly dependent on the accuracy of flashbulb memories.

In order to ensure a precise measurement of memory accuracy, it is important that the circumstances surrounding the reception of important news be documented or indexed shortly after the event occurs (Neisser, 1982). The initial documentation of the circumstances surrounding the reception event can then be compared to subsequent recalls resulting in a measure of consistency. Since the Brown and Kulik (1977) study, several researchers have compared initial documentations of an event with subsequent recalls. The level of consistency, however, varies widely. Nonetheless a trend emerges: Researchers who documented the events shortly after they occurred (Larsen, 1992; Neisser & Harsch, 1992; Weaver, 1993) have generally found less consistency than those who waited longer before documenting the event (Christianson, 1989; Conway et al., 1994; Pillemer, 1984). See Table 1 for a summary of time delays and consistency measures across various flashbulb studies.

For example, the morning after the Space Shuttle Challenger exploded in 1986, Neisser and Harsch (1992) questioned a group of participants as to how they heard the news. Over two years later, the participants' memories were tested again. Neisser and Harsch reported a high incidence of errors in the follow-up questioning. A large proportion of the participants in Neisser and Harsch's study forgot "major details", such as the informant, location, or ongoing activity. In

measuring consistency, participants were scored from 0, if no major details matched, to 7, if all major details and other information matched. Only 7% of the participants received the maximum score of 7, 9% received a score of 6, 25% were completely inconsistent, and 50% scored 2 or less on the consistency scale. Furthermore, many participants stuck to their erroneous stories; even showing them their initial reports did not aid recovery of their original memories. The errors were firmly ensconced. Other researchers have documented the reception event circumstances shortly afterwards and also found a high level of inconsistency (Larsen, 1992; McCloskey, Wible, & Cohen, 1988; Weaver, 1993).

Not all researchers report finding inconsistencies in flashbulb memories. For example, Conway et al. (1994) conducted a flashbulb memory study for hearing the news of Margaret Thatcher's resignation. The initial assessments, however, were made up to 14 days after the resignation and compared to reports collected nearly one year later. Conway et al. reported very high levels of consistency among UK participants: Approximately 59% reported the exact information at Time 2 as they had at Time 1 and 86% of the participants received a consistency score of at least 9 out of a possible 10 points (see Pillemer, 1984, for similar findings). They concluded that a special mechanism is needed to explain flashbulb memories. However, Conway et al. assumed that the initial descriptions were accurate, even though the event had occurred up to 14 days prior to the documentation of the memories.

Neisser, Winograd, Bergman, Schreiber, Palmer, & Weldon (1996) data provides a unique opportunity to examine the effect of different time delays between the event and indexing of that event on the consistency of reports. They looked at individuals' recollections regarding the Loma Prieta earthquake. Data from participants in Santa Cruz were collected 15 - 21 days after the earthquake. In contrast, data from Berkeley participants were collected only 1 - 3 days after the earthquake. According to Neisser et al., the difference in time delays may be related to the observation that the Santa Cruz participants provided more consistent accounts of their experiences than did the Berkeley students.

We consider two possible explanations for why more errors are observed when events are documented shortly after an event occurs. First, errors may be caused by the introduction of post-event information (Loftus, 1975, 1979; Loftus, Miller, & Burns, 1978; Loftus & Palmer, 1974). People continue to learn more about important public events even after they first learn of it: People hear stories from other individuals about how they received the news, they receive additional news about the event, and they may embellish their own accounts in subsequent retellings (Winningham, Orebaugh, Steves, & Weaver, 1997). The eventual story an individual settles on may be a reconstruction of information from all these sources. This reconstruction may be relatively complete within a few days, unless an individual has a reason to continue considering the event.

A second explanation is that the observed inconsistencies are an artifact of forgetting. Most information is forgotten shortly after an event (Ebbinghaus, 1885/1913). Thus the immediate recollections will be most detailed and when these memories are compared to later, more vague, recollections, many inconsistencies will be apparent. In contrast, when delayed recollections are compared to later recollections, fewer inconsistencies will be observed because both are at a similar level of vagueness.

Unfortunately, it is difficult to meaningfully compare results across studies. Researchers have measured different aspects of different events, used different scoring methods, and varied the time intervals both at the original documentation and the later recalls. Therefore, the present study was designed to examine if the time delay between the event and initial documentation is related to the consistency of reports. We examined this hypothesis by documenting participants' recollection of hearing of O.J. Simpson's acquittal either immediately (5 hours) or after a delay (1 week). All individuals provided a second recollection eight weeks after the event. We predicted that the immediate group would be less consistent between their Time 1 and Time 2 reports than the delayed group.

We decided to use O.J. Simpson's acquittal as the event in this study. The reading of the verdict was the end of a long drawn out trial, in which O.J. Simpson, an African American sports

hero, was accused of killing his ex-wife Nicole Brown Simpson and her friend Ron Goldman in 1994. This was a significant and newsworthy event. According to a 1995 CNN poll 88% of Americans learned of the verdict within 30 minutes of its announcement. Moreover, we found that over three years after the acquittal, 82% of college students surveyed would have been classified as having a flashbulb memory for this event, according to Brown and Kulik's criterion.

Method

Participants

Two groups of participants were used in this study. Each group of participants was enrolled in an undergraduate psychology course and tested as a group during class time. The two classes were both required courses for undergraduate psychology majors, therefore the classes should have been very similar. The sample was relatively homogenous, 88% of the participants were female and 97% were Caucasian. The immediate group ($n=35$) was tested five hours after the O.J. Simpson verdict was read. The delayed group ($n=30$) was tested one week after the acquittal.

Materials and Procedure

After the participants read and signed informed consent forms, the questionnaires were distributed. The questionnaires were completed during the last half of 50-minute classes. The participants were instructed to answer all questions as best as they could.

At Time 1, both groups received the same questionnaire that was designed to assess information regarding how they learned of O.J. Simpson's acquittal. The participants were asked to provide information related to 4 main categories of information: informant, time, place, and ongoing activity. The participants also rated their emotionality, and surprise, and how personally and societally important they thought the event was via 7-point scales (1 meant low and 7 meant high). Three types of rehearsal were examined by assessing the percentage of time the participants reported having thought about, talked about, or listened to radio or television programs concerning the event. Participants were not told that their memory would be tested a second time.

The Time 2 questionnaire included the same questions and instructions as the Time 1 questionnaire and two additional questions regarding the extent to which they had a visual image of the event and how unique the event was. These questions were also asked via 7-point scales.

Scoring

We employed a three-point rating scale for judging the consistency of the four main components. This approach is similar to coding systems used by other researchers (Conway et al., 1994; Neisser & Harsch, 1992). Responses were scored as either a 0, 1, or 2. A score of 2 was given if participants reported the same information at Time 2 as they did at Time 1. A consistency score of 1 was given if participants generally provided congruent information at Time 2, but not exactly what was reported at Time 1, or if the Time 2 response was more general or more specific than the Time 1 response. For example, a score of 1 was given if they originally reported being on the third floor of Miller Hall on the Western Washington University campus and then later reported being on campus. A score of 0 was given if the participant provided inconsistent information at Time 1 and Time 2: if they originally reported being at school when they heard the news, then they later reported being at home. Answers to the “When did it happen question” could receive a half a point depending on how close they were in reporting the time they had reported at Time 1. The four consistency scores were summed together to get a global consistency score, which ranged from 0-8.

To score the amount of information recalled, we counted the number of words written about the event and the number of propositions that the participants used. We quantified the propositions by counting the number of subject-verb phrases and the number of affective descriptions (e.g., surprised, angry). Subject-verb phrases indicated distinct propositions (“I was watching it on CNN”; “I was in class”; “My mom and I just looked at each other”). In some instances, the verb was understood or implied by the question. For example, in the question “Who first told you about the verdict”, writing “radio station” implies a verb.

Results

Consistency

We found a difference in the consistency of responses as a function of when the events were initially indexed, $t(63) = 2.41$, $p = .019$. The immediate group ($M = 5.00$, $SD = 2.03$) provided less consistent reports of O.J. Simpson's acquittal than the delayed group ($M = 6.20$, $SD = 1.98$). An omega-square revealed that 7.9% of the variability in the consistency was explained by the timing of the initial documentation. Thus, as predicted, we found that when memories are documented after a delay, they appear to be more consistent than memories that are documented immediately after the event. If we classify participants who had consistency scores above 7 as having flashbulb memories and those with scores below 7 as not having flashbulb memories (see Conway et al., 1994 for a similar classification criterion) then only 22.85% of the immediate group had flashbulbs, whereas 53.34% of the delay group were classified as having flashbulb memories.

Many participant ratings were reliably correlated with the consistency of reports (see Table 2). Self reported measures of emotionality, surprise, personal significance, and societal significance were not extremely high (see Table 3). In addition, we tested the possibility that the observed differences in consistency as a function of group may have been due to initial levels of the participants' appraisals. However, a series of t-tests did not reveal any difference in emotionality $t(63) = 1.38$, $p = .171$, surprise $t(63) = .518$, $p = .606$, personal significance $t(63) = 1.778$, $p = .08$, or societal significance $t(63) = .123$, $p = .902$ between the two groups. Therefore it is unlikely that the above appraisals are driving the observed effect of consistency between the groups.

How Much was Written

The effect of delay on how much the participants wrote at Time 1 and Time 2 was analyzed using 2X2 mixed model ANOVAs (delay group as a between-subjects factor and time as a within-subjects factor). For the number of words, there was no effect of group, a main effect of time, $F(1,63) = 5.31$, $p = .024$, $MSE = 103.30$, and no interaction. The participants wrote more words at Time 1 than Time 2. There was a reliable correlation in the number of words reported at Time 1 and

Time 2 for the immediate group $r = .430$ ($p = .010$) and the correlation for the delayed group approached significance $r = .341$ ($p = .065$).

For the number of propositions there was a main effect of the delay group, $F(1, 63) = 9.88$, $p = .003$, $MSE = 4.53$, a main effect of time, $F(1, 63) = 16.57$, $p < .001$, $MSE = 1.68$, and a significant interaction, $F(1, 63) = 5.37$, $p = .024$, $MSE = 1.68$. The interaction shows that although both groups included fewer propositions over time, the loss was greater for the immediate group. Put differently, although the difference between the immediate group and the 1-week delay group was apparent at both times, the difference is larger at Time 1 (see Table 4). There was a reliable correlation in the number of propositions reported at Time 1 and Time 2 for the immediate group $r = .424$ ($p = .011$) and the delayed group $r = .517$ ($p = .003$).

General Discussion

Does the timing of the initial indexing of an event influence the apparent consistency of memories for that event? Yes. We found that participants who were initially tested immediately after O.J. Simpson's acquittal provided less consistent reports than participants who were initially tested one week after the event. Our results confirm a trend observed in previous flashbulb research: Researchers who initially indexed an event shortly after it occurred (Larsen, 1992; Neisser & Harsch, 1992; Weaver, 1993) generally found lower levels of consistency than researchers who waited longer before obtaining the initial documentation event (Christianson, 1989; Conway et al., 1994; Pillemer, 1984). Therefore, it appears that the time of initial testing needs to be considered when conducting flashbulb memory studies.

There are several possible interpretations of these findings. First, this may be a simple case of "Ebbinghausian" forgetting (Ebbinghaus, 1885/1913): a rapid initial period of forgetting followed by asymptotic long-term retention. During the week immediately following O.J.'s acquittal people may have forgotten some of the details associated with the event. If so, then the delayed group's initial reports would have been less detailed than the immediate group's. If all reports become less detailed over time, then the original documentation of the delayed group

would have been more similar in level of detail to the final recollection. The immediate group would thus have a more difficult task of providing consistent reports than would the delayed group. Supporting this view, we found that more propositions were included in the immediate group's initial reports than the delayed group's.

A second explanation of the time delay effect is that errors in memory were introduced in the week immediately following the events. If this were the case, then the delayed group's errors would be included in their initial index. In contrast, the immediate group errors would have been incorporated into their memory after the initial indexing. Loftus and Kaufman (1992) noted that it is possible to rehearse correct and incorrect details of an event. If consistent event details are rehearsed there should be an increased likelihood of having a consistent report, but if inconsistent event details are rehearsed then there should be an increased chance of having an inconsistent report. Moreover, changes in memory may occur in response to the encoding of additional information that requires updating old memories. Most people talked about and saw media coverage pertaining to the O.J. acquittal during the week following those events. That additional information may have altered the original memories to the deficit of the immediate group's consistency scores.

As another way of explaining some memory errors, Brewer (1988, 1992) suggested that it is possible that some individuals recall an experience that occurred but that the experience was not the event originally described. Thus, the researchers score the Time 2 responses as incorrect since they do not match the Time 1 responses. Brewer called this the wrong time slice hypothesis (also see Neisser & Harsch, 1992). The reporting of wrong time slices is a possible source of erroneous information in flashbulb type studies. For instance, if someone initially reports that they were eating breakfast when learning of an important event, then later report that they were talking with a roommate when they heard a news broadcast about the event, the latter response would be scored as being incorrect. It is possible, however, that both events occurred, but at different times. The person may have retrieved an event associated with hearing the news, but not the originally documented event.

These explanations are not necessarily mutually exclusive, of course. Forgetting may have been more rapid in the first few days following the event. At the same time, new details -- both correct and incorrect -- may have been added to the original event memory precisely when they would have been most influential.

Brown and Kulik (1977) suggested that a special mechanism was needed in order to explain the remarkably long-lasting and detailed memories of important and consequential events. Since then, many researchers have argued about whether it is necessary to postulate a special mechanism (Conway, 1995; McCloskey et al., 1988; Neisser, 1982; 1988; Pillemer, 1990). Many of the arguments have been based on the purported accuracy (operationally defined as consistency) of so called flashbulb memories (Neisser, 1986; Thompson & Cowan, 1986). Our research indicates that interpretations of the consistency observed in various studies should consider the time interval before an initial documentation. Learning about O.J. Simpson's acquittal did not yield exceptionally high levels of surprise, emotionality, and personal significance, thus it may not have been a true flashbulb event. However, it is clear that researchers should try to obtain the initial indexing of flashbulb type events as quickly as possible. If this is not accomplished then results indicating memory consistency need to be viewed skeptically. Future research should identify the processes that are involved in undermining the consistency of reports that are immediately indexed. After an event, a memory may be gradually consolidated as people forget some information, incorporate information from other sources, and develop a narrative of the event.

References

- Brewer, W. F. (1988). Memory for randomly sampled autobiographical events. In U. Neisser & E. Winograd (Eds.), Remembering reconsidered: Ecological and traditional approaches to the study of memory (pp. 21-90). New York: Cambridge University Press.
- Brewer, W. F. (1992). The theoretical and empirical status of the flashbulb memory hypothesis. In E. Winograd & U. Neisser (Eds.), Affect and accuracy in recall: Studies of "flashbulb" memories (pp. 274-305). New York: Cambridge University Press.
- Brown, R., & Kulik, J. (1977). Flashbulb memories. Cognition, *5*, 73-99.
- Christianson, S. (1989). Flashbulb memories: Special, but not so special. Memory & Cognition, *17*, 435-443.
- CNN (1995). CNN - O.J. Simpson verdict poll - Oct 4, 1995. Available: <http://www.cnn.com/US/OJ/daily/9510/10-04/poll/ojpoll5.html>.
- Conway, M. A. (1995). Flashbulb memories. Hove, UK: Lawrence Erlbaum Associates, Publishers.
- Conway, M. A., Anderson, S. J., Larsen, S. F., Donnelly, C. M., McDaniel, M. A., McClelland, A. G. R., Rawles, R. E., & Logie, R. H. (1994). The formation of flashbulb memories. Memory & Cognition, *22*, 326-343.
- Ebbinghaus, H. (1885/1913). Memory: A contribution to experimental psychology (H. A. Ruger & C. E. Bussenius, Trans.). New York: Columbia University, Teachers College.
- Larsen, S. F. (1992). Potential flashbulbs: Memory 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.
- Loftus, E. F. (1975). Leading questions and the eyewitness report. Cognitive Psychology, *7*, 560-572.
- Loftus, E. F. (1979). Eyewitness testimony. Cambridge, MA: Harvard University Press.
- Loftus, E. F., & Kaufman, L. (1992). Why do traumatic experiences sometimes produce good memory (flashbulbs) and sometimes no memory (repression)? In E. Winograd & U.

Neisser (Eds.), Affect and accuracy in recall: Studies of "flashbulb" memories. (pp. 212-223). New York: Cambridge University Press.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. Journal of Experimental Psychology: Human Learning and Memory, 4, 19-31.

Loftus, E. F., & Palmer, J. P. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory, Journal of Verbal Learning and Behavior, 13, 585-589.

McCloskey, M., Wible, C. G., & Cohen, N. J. (1988). Is there a special flashbulb mechanism? Journal of Experimental Psychology: General, 114, 3-18.

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

Pillemer, D. B. (1984). Flashbulb memories of the assassination attempt on President Reagan. Cognition, 16, 63-80.

Pillemer, D. B. (1990). Clarifying the flashbulb memory concept: A comment on McCloskey, Wible, and Cohen (1988). Journal of Experimental Psychology: General, 119, 92-96.

Thompson, C. P., & Cowan, T. (1986). Flashbulb memories: A nicer interpretation of a Neisser recollection. Cognition, 22, 199-200.

Weaver, C. A., III. (1993). Do you need a "flash" to create a flashbulb memory?. Journal of Experimental Psychology: General, 122, 39-46.

Winningham, R. G., Orebaugh, B., Steves, R., & Weaver, C. A., III (April, 1997).

The effects of embellished rehearsals on the consistency of autobiographical memory.

Poster presented at the 43rd annual Southwestern Psychological Association.

Table 1

Summary of Time Delays and Consistency Measures for Flashbulb Studies

Study	Event	Time To First Recording	Interval	Unit of Measurement	Accuracy
Neisser et al. (1996)	News of Earthquake 2 & 9 days at Emory		1.5 years	Combined Accuracy Percentage	55%
Neisser et al. (1996)	Earthquake at Berkeley (mild)	1-3 days	1.5 years	Combined Accuracy Percentage	96%
Neisser et al. (1996)	Earthquake at Santa Cruz (severe)	15-21 days	1.5 years	Combined Accuracy Percentage	99%
Neisser et al. (1996)	News of Bridge at Berkeley	1-3 days	1.5 years	Combined Accuracy Percentage	87%
Neisser et al. (1996)	News of Bridge at Santa Cruz	15-21 days	1.5 years	Combined Accuracy Percentage	93%
Neisser and Harsch (1992)	Challenger	1 day	2.5 years	Average Weighted Accuracy Score	2.95 / 7 points (42%)
McCloskey et al. (1988)	Challenger	about 3 days	9 months	% of inconsistent responses	8.4%
	Challenger	about 3 days	9 months	% of responses not remembered	5.6%
	Challenger	about 3 days	9 months	% of more general responses	18.7%
Weaver (1993)	Gulf War	2 days	3 months	lenient scoring method	72.5% consistent
Weaver (1993)	Gulf War	2 days	1 year	lenient scoring method	71.7%
Weaver (1993)	Gulf War	2 days	1 year	strict scoring method	30.17%
Pillemer (1984)	Reagan being shot	1 month	5.5 months	consistency ratings from 1 - 6	average rating 4.88 (81%)
Conway et al. (1994)	Thatcher's Resignation for U.K. Subjects	Within 14 days	11 months	Accuracy Score (0-1)	Approximately .9
Conway et al. (1994)	Thatcher's Resignation for non-U.K. Subjects	Within 14 days	11 months	Accuracy Score (0-1)	Approximately .6

Note. Accuracy scores for Weaver's (1993) study were computed by averaging 5 canonical categories. Data from Conway et al.'s study was estimated from Figure 1 (p. 331).

Table 2

Correlations and Probabilities of Significant Predictors of Consistency (n = 65)

Variable	r	p
Personal Importance	.35	.005
Time Thinking		.26 .036
Surprise *	.25	.049
Personal Importance *	.36	.003
Societal Importance *	.28	.022
Time Thinking *	.25	.049
Visual Image *		.28 .023
Confidence *	.38	.002
Uniqueness *	.31	.011

Note. * = Time 2 responses.

Table 3

Means and Standard Deviations for Ratings of Surprise, Emotionality, Personal Importance and Societal Importance

	Emotionality	Surprise	Personal Importance	Societal Importance
<u>Time 1</u>				
Immediate Group	3.31 (1.75)	3.97 (1.77)	2.77 (1.59)	5.26 (1.52)
Delay Group	3.90 (1.65)	4.20 (1.77)	3.47 (1.55)	5.30 (1.24)
<u>Time 2</u>				
Immediate Group	3.34 (1.53)	3.89 (1.81)	2.49 (1.31)	4.83 (1.37)
Delay Group	3.73 (1.29)	4.33 (1.65)	3.13 (1.55)	5.13 (1.43)

Table 4

Means and Standard Deviations for the Number of Propositions as a Function of Group and Time

Delay Group	Time	
	Time 1	Time 2
Immediate	7.97 (1.77)	6.51 (1.82)
Delay	6.27 (1.51)	5.87 (1.91)

Figure Caption

Figure 1. The Distribution of Consistency Scores by Group.

