Journal of Experimental Psychology: Applied
2006, Vol. 12, No. 3, 142–154

1076-898X/06/$12.00 DOI: 10.1037/1076-898X.12.3.142

The Impact of Stress on Mothers' Memory of a Natural Disaster

Janat Fraser Parker and Lorraine E. Bahrick
Florida International University

Robyn Fivush
Emory University

Paulette Johnson
Florida International University

The effects of stress on mothers' recall for a major hurricane were studied. Stress was objectively defined as low, moderate, or high according to the severity of home damage. This study of 96 mothers was conducted concurrently with L. E. Bahrick, J. F. Parker, R. Fivush, and M. Levitt (1998), allowing the authors to compare child and adult recall as a function of the same stressor. There was a quadratic relationship between storm severity and total recall for adults, similar to their children. Mothers' recall increased from low to moderate severity, but recall at moderate severity did not differ from high severity. These findings help clarify the effects of stress on the amount and type of information adults recall in retrospective accounts of naturalistic, temporally extended events.

Keywords: stress and memory, memory of a natural disaster

The present study investigated mothers' recall of events surrounding Hurricane Andrew, a major Class 5 hurricane with sustained winds of 175 mph (281.64 kph). This hurricane devastated much of Southern Dade County, Florida, on August 24, 1992, and impacted the lives of families for months afterwards. Such a natural disaster allowed us to explore the effects of stress on memories of actual traumatic events and to address some of the critical issues that have been raised by literature on stress and memory. Further, we were able to compare the recall of the hurricane by adult mothers with the recall of their preschool children reported previously (Bahrick, Parker, Fivush, & Levitt, 1998). Bahrick et al. (1998) observed a quadratic function, in which moderate levels of stress were associated with the highest level of recall. We were interested in whether their mothers would show a similar relationship of stress with memory. This developmental comparison is particularly interesting, as both the children and their mothers were exposed to similar stressful events.

Research has yielded a wide variety of results on the impact of negative emotional events on memory. In a recent meta-analytic review of the effects of high stress on eyewitness memory, Deffenbacher, Bornstein, Penrod, and McGorty (2004) found support for the hypothesis that high stress negatively impacts recall (d = –.31). In his review, Christianson (1992) emphasized that the effects of stress are dependent on the type of event (e.g., real-life traumatic vs. simulated), the type of detail information (central vs. peripheral), the time of test, and the type of retrieval (recognition vs. free recall etc.). An examination of the literature suggests that actual traumatic events tend to result in high levels of recall (e.g., R. Brown & Kulik, 1977; Christianson & Hubinette, 1993; Waganaar & Groeneweg, 1990; Yuille & Cutshall, 1986, 1989; Yuille & Tollestrup, 1992), whereas simulated laboratory studies give mixed patterns often reporting lower levels of memory as stress increases (e.g., Clifford & Hollin, 1981; Clifford & Scott, 1978; Loftus & Burns, 1982).

Studies of the impact of stress on memory for real-life events are considered to be more ecologically valid and are characterized by two approaches. One approach has been to examine the recall of a shocking newsworthy event such as the death of John F. Kennedy (R. Brown & Kulik, 1977), the explosion of the Challenger shuttle (Neisser & Harsch, 1992), and the September 11, 2001 (9/11) terrorist attack (see Pezdek, 2003b) in flashbulb memory (FBM) studies. Most of this research has emphasized the recall of the circumstances surrounding learning of the event, and the results have been characterized by high levels of recall even after extended periods of time. Although some researchers (R. Brown & Kulik, 1977; Conway, 1995) attributed this high level of recall to special mechanisms, many now place FBM recall on a continuum along with recall of other less stressful material.

It should be pointed out that although these public events may be very surprising and emotional, they may not actually be traumatic to the individual reporting the event. Individuals may have only witnessed the event on television or through some other medium and may not actually have been present during the trauma. Byrne, Hyman, and Scott (2001) claimed that these events often do

Janat Fraser Parker and Lorraine E. Bahrick, Department of Psychology, Florida International University; Robyn Fivush, Department of Psychology, Emory University; Paulette Johnson, Department of Statistics, Florida International University.

This research was supported by National Science Foundation Grant DBS 9300956 to Janat Fraser Parker and Lorraine E. Bahrick.

We thank Maria Borras, Mi Yung Chung, Ana D'Costa, Giovanna Delgado, Gillian Dixon, Angela Escobar, Rita Estorino, Ivan Grossman, Juan Lata, Brenda Lundy, Millie Munoz, Judy Ramsingh, Rita Soza, and Adina Sved for their assistance in interviewing, transcribing, coding, and tabulating data; Brenda Lundy for coordinating subject recruitment and data coding; and Jim Jaccard, Mary Levitt, Bradley McAuliff, and Stacy West for their assistance in data analyses and/or graphics.

Correspondence concerning this article should be addressed to Janat Fraser Parker, Department of Psychology, Florida International University, University Park, Miami, FL 33199 E-mail: parkerjf@fiu.edu

not fit the definition of trauma in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychological Association, 1994).

Neisser et al. (1996) looked at the effect of the Loma Prieta, California, earthquake on the FBM of three different groups of participants—two from California (one group much closer to the earthquake's epicenter) and one from Atlanta, Georgia. They found that recall of the California groups was virtually perfect, whereas the Atlanta group showed significantly less recall. However, there was no correlation between stress and recall. Of course, this does not preclude the possibility of a quadratic relationship of memory with stress, although none was tested. As the average level of stress even for the California groups was not high, Neisser et al. claimed that, on average, the earthquake was not considered a traumatic event for these participants. Thus, it could well be that a restricted range of stress was sampled that precluded finding any significant relationship between stress and recall. Neisser et al. suggested that the opportunity for rehearsal of the earthquake narrative might have been an important factor in the higher level of recall for the West coast participants over the Atlanta participants, rather than greater stress levels.

There has been a relative silence regarding the effects of such disasters on the recall of the event per se rather than the circumstances surrounding learning of the event. Recently, Pezdek (2003a) tested an extension of Neisser et al.'s (1996) narrative hypothesis, namely, that it is the synergy of arousal and rehearsal that affects memory. She compared the recall of the 9/11 disaster event itself with the autobiographical aspects of the event (also known as FBM) across three different geographical areas—New York, California, and Hawaii. Her hypothesis was that the stressful event and people's autobiographical experience of the event would be processed separately, resulting in separate memories. She further claimed that the most vivid and most emotionally charged of these memories would receive the greatest rehearsal and hence be better recalled. Further, she hypothesized that event memories would be more enhanced and autobiographical memories less enhanced by the emotionality of the situation, the closer the individual was to the disaster. Consistent with her predictions, the New York sample was the most accurate on questions of event memory (large effect sizes ranging from $\eta^2 = .78$ to .96) but provided the fewest details on questions of autobiographical memory (large effect sizes ranging from $\eta^2 = .96$ to .99). Smith, Bibi, and Sheard (2003) partially supported these findings by showing that Canadians (far from the disaster) showed a statistically significant greater consistency of recall for autobiographical information than event information (effect size [*ES*] = .83; see explanation of effect size calculations in Results and Discussion). However, significant Pearson product–moment correlations between self-reported arousal level and amount of autobiographical recall were not observed; in contrast, statistically significant Pearson product–moment correlations between arousal level and event memory were observed ($r = .32$, $p < .01$). Neisser (2003), however, has criticized these studies, claiming that the greater event memory for those closer to the event may not be a function of the synergy of arousal and rehearsal but rather a function of greater exposure to the event and hence greater rehearsal.

There are flaws inherent in this approach to measuring the effects of stress on memory. Typically, there is no neutral event with which to compare the traumatic event. In one FBM study, Christianson (1989) attempted to include a control group but pointed out the lack of compatibility in the choice of the neutral event and the emotional event as well as differences in the method of memory measurement across groups. Larsen (1992) was somewhat more successful in his comparison of neutral news and personal events with upsetting news and personal events. His analysis of the memory of such events collected in his personal diaries suggests memory for the news itself improves with increases in emotional value, but memory of the surrounding context deteriorates.

The second approach of researchers interested in recall of real-life traumatic events has been to focus on recall of the stressful event itself by participants who directly experienced the event (e.g., Christianson & Hubinette, 1993; Christianson & Nilsson, 1989; Waganaar & Groeneweg, 1990; Yuille & Cutshall, 1986, 1989). Typically, these studies also show high levels of recall of the traumatic event, but they also are often devoid of neutral control events. Even if control groups are included, there are myriad factors, such as different levels of interest and general knowledge about one event over another and a priori differences inherent in the tasks themselves, which make it difficult to equate stress and control groups in every way, except for stress level.

To address this baseline issue, Yuille and Cutshall (1986) examined witnesses exposed to the same shooting event and categorized them according to the individual level of stress they experienced, with the least anxious selected as control witnesses. They observed higher levels of recall for those witnesses who reported high stress levels versus those who reported lower stress levels, but the number of witnesses was minimal, 7 in the high stress group and 6 in the control group. Further, their procedure permitted emotional stress to be confounded with other variables. For example, witnesses may have differed in access to critical information during the crime, with the more stressed individuals being closer to the central aspects of the crime and the less stressed more in the periphery (see also Christianson, 1992; Goodman & Loftus, 1989). Further, in these studies, a fairly narrow range of stress is typically sampled, restricting the likelihood of finding an effect. Finally, some field studies have even found a negative relationship between memory and stress (e.g., Byrne et al., 2001). Relying on female undergraduates' ratings on the Memory Characteristics Questionnaire (Johnson, Foley, Suengas, & Raye, 1988), Byrne et al. (2001) assessed the similarities and differences among traumatic, negative, and positive experiences. The students' ratings suggested that they were less able to recall aspects of the traumatic and negative experiences than the positive experiences (*ES*s = .10 and .07, respectively).

To overcome some of the drawbacks of field studies, researchers have carried out laboratory simulation studies. The major advantage of such studies is that accuracy could be objectively determined. Further, control events could be developed to compare with the simulated emotional event. However, even in the laboratory, it is difficult to obtain a neutral and a stressful event that are comparable on all other dimensions. Furthermore, simulated emotional events may not be able to produce the same level of stress as real-life traumatic events.

In contrast with results of the field studies, most lab studies show an impairment in memory when participants are exposed to slides (Christianson, 1984), videotapes (Clifford & Hollin, 1981; Clifford & Scott, 1978; Loftus & Burns, 1982), or live staged

scenarios (Leippe, Wells, & Ostrom, 1978) depicting an emotional incident compared with a nonemotional incident. Several researchers (e.g., Christianson, 1992; Christianson & Hubinette, 1993; Christianson & Loftus, 1991; Heuer & Reisberg, 1992; Kebeck & Lohaus, 1986) have attempted to reconcile the differences observed between and within the laboratory studies and field studies by examining the relation between the type of information (central vs. peripheral) and the level of recall as a function of emotion (see Christianson, 1992; Heuer & Reisberg, 1992). Reisberg and Heuer (in press), in their recent analysis of the influence of emotion on memory, pointed out that laboratory simulation studies (J. M. Brown, 2003; Burke, Heuer, & Reisberg, 1992; Christianson, 1984; Christianson & Loftus, 1987, 1991) have shown that central information is recalled better in emotional than nonemotional situations (e.g., *ES* = .14 in Burke et al., 1992) but peripheral information is recalled better in nonemotional situations than emotional situations (e.g., *ES* = .06 in Burke et al., 1992). Central material can be defined in numerous ways such as plot relevance as well as perceptual–spatial location. When the results of several studies were examined from this perspective, there was a pattern of positive effects of stress on central material but negative or no effects on peripheral material (e.g., Clifford & Hollin, 1981; Loftus & Burns, 1982). Further Steblay's (1992) meta-analysis of the weapon focus effect is consistent with these findings. The expectation was that a weapon (central) would decrease the ability of the eyewitness to adequately encode and later recall peripheral details (e.g., perpetrator). Comparisons between weapon-present and weapon-absent conditions established that descriptions of the perpetrator were more accurate in the weapon-absent conditions (*d* = .55).

Similar dissections of field and FBM studies also suggest that emotion facilitates recall of central material while undermining recall of peripheral material (e.g., Christianson & Hubinette, 1993; Larsen, 1992). These findings across laboratory and field studies are consistent with Easterbrook's (1959) cue utilization hypothesis that claims there is a progressive restriction in the range of cues used as a function of increases in arousal, resulting in a "narrowing" of attention. This focusing of attention is presumed to result in greater memory for central items and a reduction in memory for peripheral items. For example, Safer, Christianson, Autry, and Osterlund (1998) showed that "tunnel memory" occurred in memory for traumatic slides (e.g., a woman with her neck slashed, a man holding a bloody knife) versus neutral slides (*ES* = .03). Further, they claimed that tunnel memory involves processing emotion-arousing information more elaborately and inhibiting boundary extension.

There are obvious methodological differences across studies as well. We have discussed the difficulty in attempts to match stress levels across laboratory and field studies. Further, numerous stress measurement techniques (subjective, behavioral, physiological) have been used, rendering it difficult to operationalize stress and to equate stress ratings across studies. Likewise, memory has been tested in a variety of ways for many aspects of the events. Variations of recognition tests, recall tests, and Memory Characteristics Questionnaire ratings (Byrne et al., 2001) have been used across studies.

To resolve these conflicting findings, Deffenbacher (1983) initially hypothesized an inverted-*U*-shaped relation between memory and stress, consistent with the Yerkes–Dodson (1908) law. He applied the Easterbrook (1959) attentional narrowing hypothesis to the existing literature and suggested that studies with low levels of stress did not benefit from focused attention, studies with moderate levels of stress were positively impacted, and studies with high levels of stress were negatively impacted because the attentional narrowing resulted in too dramatic a confining of information. More recently, Deffenbacher (1994) and others (Christianson, 1992; Deffenbacher et al., 2004; Reisberg & Heuer, in press) have shifted their attention from the Yerkes–Dodson proposal, citing its difficulty in handling different types of memory and different types of emotion and the lack of any direct evidence of the inverted-*U*-shaped function relating stress and memory. Rather, Deffenbacher (1994) and Deffenbacher et al. (2004) provided an integrative theoretical alternative to such a unidimensional arousal theory. Their model predicts that high levels of cognitive anxiety and physiological activation will result in a drop in performance, and their predictions were supported by a meta-analysis of both face identification and eyewitness recall of details (*d* = –.31). Regardless of the underlying mechanisms, both approaches predict a drop in recall at high levels of stress. In order to test for an inverted-*U*-shaped function, however, one needs at least three levels of stress, and existing studies typically include only one or two levels. Bahrick et al. (1998) addressed this issue by conducting a naturalistic study of the relationship between memory and stress in which three levels of stress were sampled. We found an inverted-*U*-shaped curve for young children's memory of Hurricane Andrew as a function of stress.

Assessing memory for Hurricane Andrew allowed us to resolve a number of methodological difficulties. It was a real-life, traumatic event that elicited both a high degree and a wide range of stress. It also provided a natural low-stress control group (those who prepared for the hurricane but experienced only a storm), allowing us to compare recall across three levels of stress for a single event. Stress was objectively defined as low, moderate, or high according to the severity of damage to the mothers' homes, allowing us to compare recall of the same event across three levels of stress.

Prior research on the effects of natural disasters or trauma on adults has focused largely on socioemotional issues such as coping strategies after the Loma Prieta earthquake and the Persian Gulf War (Pennebaker & Harber, 1993), posttraumatic stress disorder after a school sniper attack (Pynoos et al., 1987) or after the Mount St. Helens's volcano eruption (Shore, Tatum, & Vollmer, 1986), or nightmares after the Loma Prieta earthquake (Wood, Bootzin, Rosenhan, Nolen-Hoeksema, & Jourden, 1992). There has been a dearth of research on the memory of natural disasters by individuals who experienced the event.

A recent study by Ackil, Van Abbema, and Bauer (2003) has compared children's recall of a tornado with nontraumatic events (most were positive experiences) in the context of collaborative mother–child interactions. They found that mother–child recollections of the tornado were longer, more narratively coherent, and more complete than their recollections of benign events. However, this study still has the inherent drawback of comparing a traumatic event with a nontraumatic event that may differ on many dimensions.

Although Ackil et al. (2003), Neisser et al. (1996), Pezdek (2003a), and Smith et al. (2003) have advanced understanding of memory for disasters as a function of stress, the present study was

able to investigate several additional aspects that were not addressed by these studies. Neisser et al. examined memory for the same earthquake with three different groups of people but mainly focused on recall of the FBM components of the event. Some researchers who investigated the 9/11 disaster (e.g., Pezdek, 2003a; Smith et al., 2003) looked at both FBM and the memory of the event itself. However, there have not been any reports of the memory of individuals who were personally involved in the 9/11 disaster.

In the current study, we were particularly interested in assessing memory for the emotional event itself (the hurricane, the hurricane preparation, and the aftermath) as a function of stress experienced by the mothers. We examined more than two levels of stress so that the inverted-*U*-shaped function predicted by the Yerkes–Dodson (1908) law could be tested. Ours was the first research to explore the effects of more than two levels of stress on adults' memory of a natural disaster. The present study was conducted concurrently with the Bahrick et al. (1998) study of children's recall of Hurricane Andrew. We assessed the memory of the children's mothers in a similar open-ended interview procedure with free recall followed by prompted recall. Thus, this study permits a developmental comparison of memory for a natural disaster with the same measures taken for the same event for the same time periods.

We focused on several primary research questions. Our first question addressed whether mothers' recall would be related to stress by an inverted-*U*-shaped function. Thus, we compared recall at the low, moderate, and high storm severity levels, with the expectation that the highest level of recall would be observed for the moderate severity level. Second, we investigated whether mothers recalled more information about the hurricane itself, the hurricane preparations, or the hurricane aftermath and whether recall of these time periods was differentially affected by hurricane severity level. Third, once free recall of an event was completed, we expected that mothers, like their children, would show a great deal of additional recall following specific prompts. We also explored whether there would be differences in the type of recall as a function of stress. Fourth, we hypothesized that recall of actions would be greater than descriptions, which in turn would be greater than recall of internal states, consistent with the child study. Further, we explored whether the recall of these three types of content would differ as a function of hurricane severity level. Fifth, in order to address whether severity level impacted central versus peripheral information differentially, we compared recall of central and peripheral information across severity levels. Finally, we wished to determine whether other factors, such as socioeconomic status (SES), rehearsal, or retention interval, could account for the observed results.

Method

Participants

Ninety-six mothers (age: $M = 34.6$ years, $SD = 4.4$ years) and their children aged 3 and 4 years were recruited for participation through local preschools. There were 4 more children than mothers because one set of triplets and two sets of twins were tested. Only the data of the mothers are reported here, as the children's data have been published previously (Bahrick et al., 1998). Mothers were recruited from an area spanning a 30-mile (48.28-km) distance in Miami-Dade County, Florida, to include low ($n = 17$), moderate ($n = 55$), and severe ($n = 24$) hurricane exposure (see *Storm Severity Classification* section for details). Mothers were primarily from middle-class families, with family incomes averaging between $40,000 and $60,000, and an average of 15.4 ($SD = 2.6$) years of education. The ethnic background of the sample was 70.2% White non-Hispanic, 19.1% Hispanic, 3.2% African American, and 7.5% other. All families weathered the storm in their own homes.

Interview

The interview took place in a single session and comprised four parts: questions about the storm experience in general, the prehurricane preparations (Time Period 1 [T1]), the storm itself (Time Period 2 [T2]), and the posthurricane recovery (Time Period 3 [T3]). This interview was the same as that presented to their children (see Bahrick et al., 1998, for details). The interviewer initially asked an open-ended question about the general storm experience to elicit as much free recall as possible. This was followed by nondirective prompts, such as "What else?" and "Tell me more." Next, the interviewer elicited memory about each of the three time periods separately. Presentation order for each of these parts was counterbalanced across participants. Each part began with a general open-ended question and was followed by a standardized series of increasingly more specific questions (e.g., "What did you and your family do to get ready *before* the hurricane came?"). When nondirective prompts failed to elicit more information, category prompts were given (e.g., "What did you do inside–outside the house to get ready for the hurricane?"). This was followed by specific item prompts (e.g., "Did you do anything to the windows? What?"). The interview took approximately 45 min.

Questionnaire Information

Mothers were given a sum of $10 to complete and return to the laboratory a packet of nine questionnaires. These questionnaires included two measures of child behavior, the Child Well-Being Survey and the Child Frederick Reaction Index (see Bahrick et al., 1998), and six measures designed to assess the mother's stress and changes in her behavior as a function of the hurricane. Results on these measures will not be included in this article.

Procedure

Testing took place an average of 114 days ($SD = 28.7$) after the hurricane, primarily in 23 preschools in Miami-Dade County, where the mothers' children attended, or in a few cases in public locations such as libraries or churches, chosen so that the environment would be neutral with respect to retrieval cues (homes of many mothers were still damaged). Seven different trained interviewers (all women between the ages of 18 and 35) each tested approximately one seventh of the mothers. Interviews were tape-recorded and then transcribed verbatim.

Immediately after the interview, each mother filled out a hurricane severity single-item rating scale. It was designed to objectify the degree of storm exposure and served as a basis for classifying the mothers into high, moderate, or low storm severity groups. Mothers also completed a subjective stress scale, a questionnaire assessing demographic information, the extent of damage to their homes, the nature and duration of interruption in basic services, and rehearsal of hurricane-related events for each time period. The subjective stress scale required the mothers to rate on a 4-point scale, ranging from 1 (*extremely frightened and upset*) to 4 (*extremely relaxed and happy*), how they felt during each of the three periods of the hurricane. The rehearsal questionnaire required the mother to rate how often the family talked about the hurricane in the presence of the child during the first week after the hurricane, the most recent week, and the period in between. The ratings were made on a 3-point scale, ranging from 0 (*none*) to 1 (*once a day*) or 2 (*several times a day*), and were broken down

according to time periods (T1, T2, T3). A composite score was obtained by averaging across these time periods.

Memory Interviews

Memory interviews were coded according to a propositional analysis that took place in four stages (see Bahrick et al., 1998, for details).

Coding for Propositional Units and Hurricane Relevance

All information was broken down into propositional units. A propositional unit was defined as a clause that contained a subject and a verb, either explicit or implied. Next, the propositional units were scored as to their relevance to the hurricane event. This was the primary dependent variable. All information about and related to the hurricane, getting ready for it, and the aftermath period were judged relevant. We also judged as relevant general knowledge, opinions, and background information that helped in understanding the hurricane events ("We went over to see my mother in Sarasota" "It was a nursery growing bromeliads and orchids" "We just called it the bad storm or the hurricane"). Propositions were deemed irrelevant under the following categories: off topic talk ("Excuse me one second"), conversational fillers ("so er really, you know"), false starts ("I mean . . . he may have followed . . ."), questioning the experimenter ("When we saw all the damage?"), uncodable information (cannot understand what the mother is saying), refusals ("That's all"), repetitions (the mother repeats in gist a prior statement), confirmation–denials (e.g., "yes," "no"), and statements corrected later.

Coding for Content

After coding for propositional units, we coded all hurricane-relevant propositions for content according to three major categories: actions, descriptions, or internal states. Each proposition was also coded as to whether it was relevant to hurricane preparations, damage, or repair (PDR).

Three major content categories. An *action* was defined as a proposition that had a clear agent of action. This included physical activity or motion, direct or implied, and any negation of activity as long as the agent of action was the subject (e.g., "We boarded up the windows with the mattresses" "I guess their equipment blew off"). Most other propositions without a clear agent of action were coded as *descriptions.* These were defined as propositions that mentioned what things looked like or appeared to be and did not specify an agent. Use of verbs such as *be*, *seem*, *appear*, *have*, and so forth typically characterized descriptions ("We had a gas grill" "They were a fire hazard"). Propositions were coded as *internal states* if they expressed emotion or affect ("scared," "happy," "laughing," "crying," "glad") or if they mentioned the following internal states: "think," "hungry," "tired," "know," "dream," "lucky," "remember," "want," and "understand" ("He really didn't understand").

PDR content. Each proposition was further categorized as to whether it described preparation ("The kids helped us tape up the windows"), damage ("We heard breaking glass"), or repair–cleanup ("My husband had to chop the trees"), or whether it was not related to these activities.

Coding for Time Period and Type of Recall

Time period. Each proposition was rated as to whether the content was relevant to T1, the hurricane preparations; T2, the hurricane itself; T3, the hurricane recovery; or whether it was unrelated to a given time period, T4. Occasionally, mothers supplied relevant information that was not specific to a time period, such as facts, opinions, or attitudes, and these were classified as T4 (e.g., "The hurricane was called Andrew"). Often the coder had to use context to judge the time period.

Free versus prompted recall. The interview was structured to facilitate investigation of free versus prompted recall because it began with a general open-ended question and was followed by increasingly more specific prompts. Propositions were thus coded according to the specificity of the interviewer's question that elicited the proposition. All propositions elicited by the general open-ended question and the open-ended questions that began recall from each time period were classified as free recall, and all propositions elicited by category and specific item prompts thereafter were classified as prompted recall.

Interrater Reliability

Interrater reliability was established between the primary coder and a secondary coder for all five measures coded (propositional units, major content, PDR content, time period, and specificity). For three of the measures—the number of propositional units, time period, and specificity—new coders were trained, and intercoder reliability was established between the primary and secondary coder for 13–17 transcripts (representing approximately an equal number of mothers of low, moderate, and high storm exposure) on one quartile (randomly selected) of each transcript. Interrater reliability was 97% for propositional units, 93% for time period, and 93% for specificity. For the remaining two measures—major content and PDR content—the primary coder was also responsible for coding all of the transcripts of the children at the same time (Bahrick et al., 1998). Because major content and PDR content were coded in a manner identical to that of the children, interrater reliabilities for the children's transcripts between the primary and secondary coder were used. They averaged 95% for major content and 97% for PDR content.

Results and Discussion

The major analyses explored how the mothers' recall about the hurricane was influenced by stress as defined by high, moderate, or low storm severity. Further, the total number of propositions generated were analyzed as a function of time period (preparations, hurricane itself, aftermath), type of recall (free vs. prompted), type of major content (actions, descriptions, internal states), PDR content, as well as storm severity. Finally, the effects of factors including rehearsal, retention interval, family income, home value, and mothers' education on the main findings were examined.

Throughout this article, we report an unbiased effect size measure that estimates the proportion of explained variance. The estimator is based on epsilon squared (Jaccard, 1998, 2002). In cases in which epsilon squared was computed in the context of a multifactor design, this measure represents a partial epsilon squared. For repeated measures factors, epsilon squared was calculated relative to the within-group error term. For between-subjects factors, epsilon squared was calculated with the total error term. If an effect size estimate yielded a negative value, it was set to zero. We will use the generic term *ES* to refer to this standardized effect size and common standards for interpreting this effect size set .01 as small, .06 as medium, and .14 as large.

Storm Severity Classification

Mothers were classified into one of three stress levels (high, moderate, or low storm severity) depending on their responses to a questionnaire regarding storm exposure. They were instructed to read seven descriptions of storm damage and to circle which description best described the effect of Hurricane Andrew on their family. If the storm penetrated the perimeter of the home while the family was inside and the home was initially uninhabitable (scores of 1 and 2), a rating of high severity was given. These occupants

experienced breaking glass, flying objects, and physical danger. A rating of moderate severity was given if much property damage occurred during the storm but the home was left habitable (scores of 3, 4, and 5). Although the occupants often feared physical harm, there was little actual physical danger in the home. Much of the interior property damage occurred after the storm as water continued to leak in through holes in the roof and so forth. Both the high and moderate severity groups experienced extensive clean-up periods following the storm. Finally, a rating of low severity was given if families experienced little or no damage to their homes and at most found debris in their yards (scores of 6 and 7).[1] Although these ratings of stress are ordered by objective criteria, they are also relative to the range of stress experienced by the mothers in this study, and thus, it cannot be determined whether moderate stress in our study is equivalent to high, moderate, or low stress in another study. There were 17, 55, and 24 mothers who fell into the low-, moderate-, and high-severity conditions, respectively. The average value of the home did not differ as a function of hurricane severity, $F(2, 75) = 0.94$, $MSE = 9{,}245{,}476{,}794$, *ns* ($ES = 0$). Damage to home, $F(2, 74) = 58.21$, $MSE = 790{,}811{,}220$, $p < .01$ ($ES = .60$), damage to contents, $F(2, 66) = 17.69$, $MSE = 964{,}882{,}594$, $p < .01$ ($ES = .33$), days without electricity, $F(2, 83) = 11.32$, $MSE = 170$, $p < .01$ ($ES = .20$), days without running water, $F(2, 83) = 5.93$, $MSE = 53$, $p < .01$ ($ES = .10$), and days without phone service, $F(2, 81) = 9.38$, $MSE = 914$, $p < .01$ ($ES = .17$), differed significantly across groups in the direction expected, corroborating our hurricane severity measure.

Amount Recalled as a Function of Stress

The number of hurricane-relevant propositions generated by mothers ranged from 126 to 1,436, with an overall mean of 520.1 ($SD = 288.5$). The mean number of propositions recalled for low, moderate, and high severity was 400.6, 599.8, and 559.9, respectively. The distribution was skewed, with a skewness value of .87 and a standard error of .25. Because the skewness value was more than twice its standard error, log transforms were performed, and the log (base 10) of the total number of propositions generated served as our primary dependent measure.

To assess the nature of the relationship between stress and memory, we made comparisons between recall at the low-, moderate-, and high-severity levels (see Figure 1). The low-severity group recalled less than the moderate-severity group, $t(93) = 3.58$, $p < .01$ ($ES = .11$), but there was no significant difference between the moderate- and high-severity groups, $t(93) = 0.72$, *ns* ($ES = 0$). There was a moderate effect size for the quadratic trend relating the amount recalled and stress ($ES = .07$).

These findings show that recall initially increased as a function of storm severity but then remained at the same level for the moderate- and high-severity levels. The recall of the same hurricane by the children of these mothers (Bahrick et al., 1998) also showed a quadratic trend (inverted-*U*-shaped function) with storm severity. Both age groups showed an increase in recall from low to moderate severity, but children showed a significant drop in recall from moderate to high severity, whereas their mothers' recall did not differ significantly for moderate and high similarity. The children appeared to be more strongly affected at the highest level of stress.

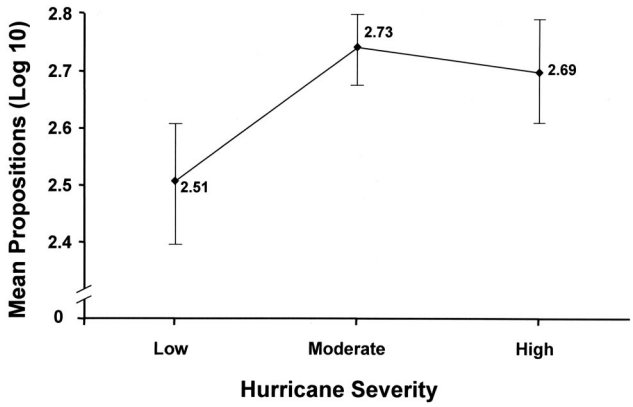

Figure 1. Mean number of propositions (log 10) recalled and 95% confidence intervals (represented by the error bars) as a function of storm severity.

To compare children with their own mothers directly, we ran a two-way analysis of variance (ANOVA), with storm severity (defined for children in the same way as the mothers) as a between-subjects factor and mother–child dyad as a within-subjects factor, which clearly showed that mothers recalled more propositions than their children, $F(1, 93) = 274.65$, $MSE = 0.04$, $p < .01$ ($ES = .74$). However, there was no statistically significant interaction of mother–child dyad with storm severity, $F(2, 93) = 0.73$, $MSE = 0.04$, *ns* ($ES = 0$). As we had sufficient power (99%) to obtain a medium effect size (population $ES = .06$), these results are not consistent with sizeable and meaningful age differences as a function of stress. However, these findings are consistent with our hypothesis that mothers' recall would be related to stress by a quadratic function, similar to their children.

Amount Recalled as a Function of Storm Severity, Time Period, and Type of Recall

More detailed analyses were conducted to examine memory as a function of time period (T1: hurricane preparations; T2: hurricane itself; T3: hurricane aftermath) and type of recall (free vs. prompted). The mean number of propositions that could be classified into one of the three time periods was 514.7 compared with a mean of 520.1 total propositions.

Table 1 reports the results of a three-way ANOVA on the log of the total number of propositions, with storm severity as a between-subjects factor and time period and type of recall as within-subjects factors. In order to determine if there were differences across the three time periods (T1, T2, and T3), we conducted tests

[1] Note that our categorization of low, moderate, and high stress was slightly different from that used in the Bahrick et al. (1998) study of children's memory for Hurricane Andrew. In particular, ratings of 1–3 were categorized as high for the children, whereas in this study, only ratings of 1–2 were categorized as high. This was because we wanted the high category to reflect only the most extreme levels of stress, because if memory declines at particularly high levels of stress, this stricter categorization scheme would provide maximal opportunity to observe a decline. Further, analyses revealed somewhat clearer differences in memory among severity ratings when this definition of stress was used.

Table 1
Analysis of Variance for Amount Recalled as a Function of Storm Severity, Time Period, and Type of Recall

Source	*df*s	*MSE*	*F*	ε^2 or partial ε^2
Between-subjects				
Storm severity	2, 93	0.32	5.99*	.10
Storm severity quadratic	1, 93	0.32	6.47*	.06
Within-subjects				
Time period	2, 186	0.08	39.74**	.29
Storm Severity × Time Period	4, 186	0.08	2.00	.02
Type of recall	1, 93	0.22	29.39**	.23
Storm Severity × Type of Recall	2, 93	0.22	0.13	.00
Time Period × Type of Recall	2, 186	0.11	12.53**	.11
Storm Severity × Time Period × Type of Recall	4, 186	0.11	1.13	.00

* $p < .05$. ** $p < .01$.

on the log of the total number of propositions. Mothers recalled significantly more propositions for T3 ($M = 262.1$) than for T1 ($M = 123.4$), $t(93) = 7.88$, $p < .01$ ($ES = .39$), or T2 ($M = 129.2$), $t(93) = 8.00$, $p < .01$ ($ES = .40$). The effect size for the difference between recall of propositions from T1 and T2 was 0.

These findings address our research question investigating whether mothers recalled more information about the hurricane itself, the hurricane preparations, or the hurricane aftermath. Clearly, mothers recalled approximately twice as many propositions about the hurricane aftermath than about the preparation period or about the hurricane itself. These differences should be viewed in the context of the nature and extent of the three temporal events. The preparation period was approximately 2 days long and quite similar across mothers of different stress levels, the hurricane itself was approximately 8 hr long, and the aftermath ranged from a few days to many weeks. Although the hurricane itself was the shortest period, mothers recalled as much about this period as the preparation period, likely reflecting an enhanced memory for an emotionally salient part of an event. For both children and adults, the amount recalled about the hurricane itself was high in relation to the amount of time elapsed.

Figure 2 displays the mean log number of propositions recalled for each time period as a function of storm severity. We explored whether recall of these time periods was differentially affected by hurricane severity level. The difference in number of propositions between low and moderate severity and between moderate and high severity was examined for all three time periods. The increase in propositions from low to moderate severity was significant for T2, $t(93) = 3.51$, $p < .01$ ($ES = .11$), and T3, $t(93) = 3.74$, $p < .01$ ($ES = .12$), but not for T1 ($ES = .01$). The change in the number of propositions from moderate to high severity was not significant across any of the three time periods, $t(93) = 0.47$, 0.24 and 0.35, *ns*, for T1, T2 and T3, respectively (all *ES*s = 0).

The relationship between storm severity and recall is similar for T2 (hurricane) and T3 (aftermath), whereas T1 (preparations) looks somewhat different from the other two periods. There is a relatively flat function for T1, whereas there is an increase in propositions from low to moderate severity, with little or no drop in recall thereafter for T2 and T3. This flat function for T1 is not surprising because mothers all experienced the same kind of preparation activities regardless of the level of hurricane severity as it was not known where the storm would make landfall. However, during T2 and T3, mothers in the low-, moderate-, and high-severity groups may have had increasingly more hurricane-related activities and thus more potentially recallable material. For example, mothers who experienced high storm severity had more distinctive events during T2, the storm itself (windows breaking and people moving from room to room to avoid flying debris etc.). Likewise, mothers who experienced high storm severity had the largest number of days without electricity, running water, or telephone and so forth in T3, the aftermath, creating unusual circumstances for a longer period of time. However, in spite of this increase in amount of potentially recallable material, recall at the highest level of severity did not increase; it dropped slightly or remained the same as for the moderate level of severity for both T2 and T3.

This inferred attenuation of recall with high storm severity is consistent with the recall of children in Bahrick et al. (1998), who

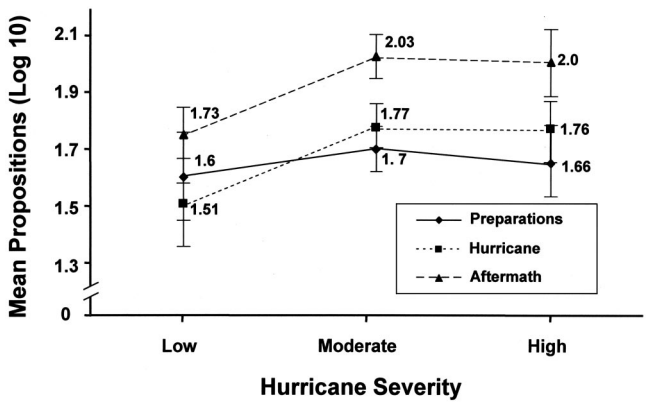

Figure 2. Mean number of propositions (log 10) recalled and 95% confidence intervals (represented by the error bars) for each time period as a function of storm severity.

showed a quadratic trend with level of storm severity. In spite of the potential for higher levels of recall at high levels of severity, the children's recall was negatively impacted, particularly for the aftermath. It appears that at high levels of storm severity, both children's and their mothers' memories were negatively affected by the stress.

The three-way ANOVA of Storm Severity × Type of Recall × Time Period (see Table 1) revealed a main effect of type of recall, with significantly more information elicited by the prompted than free-recall questions (untransformed means of 309.9 and 204.7, respectively). There was no difference in type of recall as a function of storm severity. Figure 3 displays the mean log number of propositions for free versus prompted questions as a function of time period. The superiority of prompted over free recall was significant for T1, $t(93) = 3.22$, $p < .01$ ($ES = .09$), and T3, $t(93) = 7.37$, $p < .01$ ($ES = .36$), but not significant for T2, $t(93) = 1.15$, *ns* ($ES = 0$).

Our third hypothesis was that there would be higher levels of prompted recall than free recall, in line with the results of the child study. Overall, prompting did lead to reporting of a great deal more information, 51% more additional information than reported through free recall. The interesting aspect of this comparison is not that prompted recall per se elicits more information than free recall but that so much additional information can be elicited after free recall. After mothers said they had no more information to report, specific prompts such as "What was broken?" elicited additional information. The current findings illustrate that considerably more information can be obtained about stressful events by adding specific prompts after free recall: 45% more for the preparations (free and prompted untransformed means of 50.33 and 73.07, respectively), 90% more for the aftermath (free and prompted untransformed means of 90.40 and 171.65, respectively), and 2% more for the hurricane period itself (free and prompted untransformed means of 63.98 and 65.22, respectively). Although the amount of additional prompted recall was significantly greater than free recall only for the preparation and aftermath periods, there still was a sizeable amount of additional information obtained through prompting even for the most stressful portion of the event.

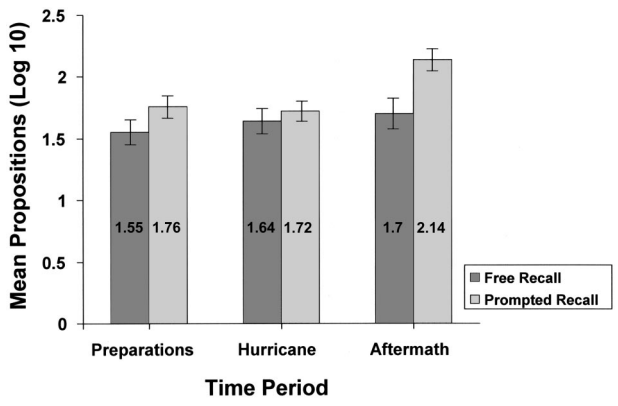

Figure 3. Mean number of propositions (log 10) and 95% confidence intervals (represented by the error bars) for free and prompted recall as a function of time period.

Memory Content as a Function of Stress

Actions, Descriptions, and Internal States

Figure 4 displays the mean number of propositions classified into each content category (actions, descriptions, internal states) as a function of storm severity. Mothers generated a mean of 242.8 actions, 208.9 descriptions, and 56.9 internal states. A two-way ANOVA, with storm severity as a between-subjects factor and content category as a within-subjects factor, was conducted on the log of the total hurricane relevant propositions. Mothers produced significantly more actions than descriptions, $t(93) = 3.56$, $p < .01$ ($ES = .11$), and more descriptions than internal states, $t(93) = 13.95$, $p < .01$ ($ES = .67$). As hypothesized, recall of actions was greater than recall of descriptions, which was greater than recall of internal states, consistent with the child study (Bahrick et al., 1998).

We explored whether the recall of the three types of content would differ as a function of hurricane severity level. Hence, each of the mother's content categories was examined separately as a function of stress. There was a significant increase in recall from low to moderate severity for actions, $t(93) = 3.23$, $p < .01$ ($ES = .09$), descriptions, $t(93) = 3.21$, $p < .01$ ($ES = .09$), and internal states, $t(93) = 2.39$, $p < .05$ ($ES = .05$). There were no significant differences between moderate and high severity for actions, $t(93) = 0.84$, *ns* ($ES = 0$), or descriptions, $t(93) = 0.12$, *ns* ($ES = 0$). However, there was a significant drop in recall from moderate to high severity for internal states, $t(93) = 2.00$, $p < .05$ ($ES = .03$). There was a moderate effect size for the quadratic trend relating the amount recalled and stress, for actions ($ES = .06$) and for internal states ($ES = .07$), and a small to moderate effect size for descriptions ($ES = .04$).

To compare mothers with their own children, we carried out two-way ANOVAs, with storm severity as a between-subjects factor and mother–child dyad as a within-subjects factor, on each of the content categories: actions, descriptions, and internal states. There was no evidence of significant interactions of severity level with age for descriptions, $F(1, 93) = 0.66$, $MSE = 0.05$, *ns* ($ES = 0$); actions, $F(1, 93) = 0.91$, $MSE = 0.05$, *ns* ($ES = 0$); or internal states, $F(1, 93) = 0.41$, $MSE = 0.08$, *ns* ($ES = 0$). As we had sufficient power (99%) to obtain medium effect sizes in all three analyses (population $ES = .06$), these results are not consistent with sizeable and meaningful age differences in content category recall as a function of storm severity. Children showed a curvilinear relation between recall and stress, with a significant drop from moderate to high severity for all content categories including actions, descriptions, and internal states (Bahrick et al., 1998). Mothers showed an increase in recall from low to moderate severity, with no change for moderate to high severity for actions and descriptions. However, there was a significant drop in recall from moderate to high severity for internal states. In light of Deffenbacher's (1994) revised three-dimensional surface relating stress and performance, mothers may be experiencing cognitive anxiety as well as high arousal when discussing emotional material related to the hurricane. As level of stress increases, these combined effects may impair recall of internal states. Children, on the other hand, may experience this high level of stress regardless of the type of content and hence show a drop in overall recall.

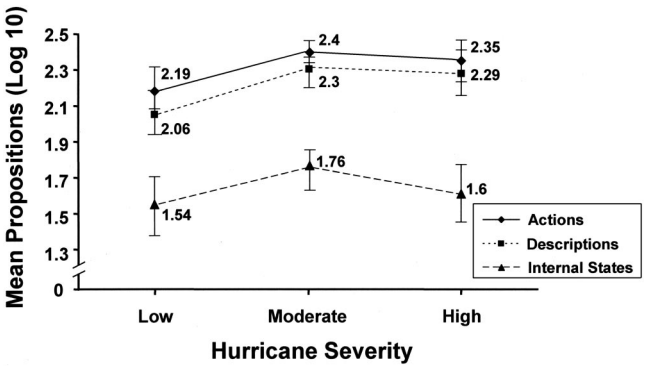

Figure 4. Mean number of propositions (log 10) recalled and 95% confidence intervals (represented by the error bars) as a function of content category and storm severity.

PDR Versus Other Content

The content of recall was also classified as related or unrelated to hurricane PDR. The PDR content was considered central, whereas the other content was considered peripheral. To address our final research question, we explored whether mothers who were more severely stressed talked less about the direct effects of the storm (PDR). A two-way ANOVA, with storm severity as a between-subjects factor and PDR content (PDR vs. other) as a within-subjects factor, was conducted on the log of the total number of hurricane propositions recalled. In addition to the main effect and the quadratic trend of storm severity previously observed, mothers showed a main effect of PDR content, $F(1, 93) = 588.36$, $MSE = 0.04$, $p < .01$ ($ES = .86$). Significantly more information was recalled about material unrelated to PDR ($M = 447.8$) than material related to PDR ($M = 72.3$) content. There was no significant interaction of PDR content with storm severity, $F(2, 93) = 0.84$, $MSE = 0.04$, *ns* ($ES = 0$), showing that stress level did not influence the amount of PDR content versus other content recalled.

Storm severity did not differentially impact the mothers' recall of PDR versus other content. Both PDR content and other content recall showed similar quadratic functions with storm severity. There was no evidence of an increasing tendency to talk about PDR content as storm severity increased. If PDR content is characterized as exclusively central information, then the present results do not show evidence of attentional narrowing with stress. Although the PDR content category was not designed a priori to be relevant to the literature on attentional narrowing and stress (Easterbrook, 1959), these findings suggest that mothers, like their 4-year-old children, are less likely to exhibit attentional narrowing with increasing stress, whereas their 3-year-old children do (see Bahrick et al., 1998).

Relations Among Recall, SES, Rehearsal, and Retention Interval

The effects observed in the above analyses were reexamined in the context of a number of covariates to determine whether they significantly qualified the main results, including any potential differences across severity groups as a function of SES (family income, mother's education, home value), rehearsal of events surrounding the hurricane, and retention interval.

SES. Table 2 depicts the mean family income, value of the family home, and number of years of education for the mother as a function of storm severity. There were significant Pearson product–moment correlations between family income and value of family home ($r = .38$, $p < .01$), between family income and number of years of mother's education ($r = .33$, $p < .01$), and between value of family home and number of years of mother's education ($r = .27$, $p < .05$). As there was substantial overlap among these variables, a composite z SES variable was calculated from the z scores of the three variables. This composite SES variable did not significantly predict amount recalled ($r = 0$), and in the analysis of covariance, $F(1, 74) = 0.00$, $MSE = 0.05$, *ns* ($ES = 0$), there was no effect. As the sample size yielded only 77% power to detect a medium effect (population $r = .30$; Cohen, 1988) between z SES and propositions recalled, we cannot with confidence claim that there is no relationship between z SES and recall. However, the main effect of storm severity, $F(2, 74) = 4.25$, $MSE = 0.05$, $p < .05$ ($ES = .08$), as well as the quadratic function relating storm severity and amount recalled $F(1, 74) = 6.81$, $MSE = 0.05$, $p < .05$ ($ES = .07$), were significant after the means were adjusted for the effect of SES.

Rehearsal. We were also interested in whether the functions relating amount recalled and stress could be mediated by differential rehearsal across the storm severity groups. Perhaps families who experienced high storm severity discussed the hurricane-related events more than families who experienced low or moder-

Table 2
Socioeconomic Status (SES): Means and Standard Deviations for Family Income, Years of Education for Mother, and Home Value as a Function of Storm Severity

	Storm severity											
	Low			Moderate			High			Total		
SES	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*N*
Family income	5.18	1.47	17	5.44	1.70	54	5.39	1.47	23	5.34	1.55	94
Mother's education (in years)	15.47	2.74	17	15.50	2.70	54	15.17	2.17	23	15.38	2.54	94
Home value ($)	159,091	59,111	11	157,678	117,760	45	124,818	47,959	22	147,198	74,943	78

Note. For income, 1 = under $10,000, 2 = $10,000–$19,999, 3 = $20,000–$29,999, 4 = $30,000–$39,999, 5 = $40,000–$59,999, 6 = $60,000–$79,999, and 7 = over $80,000.

ate storm severity. Recall that a composite score was derived on the basis of an average of the amount the family talked to the child and in front of the child about the hurricane preparations, the hurricane itself, and the aftermath. This measure is a rough indication of rehearsal for the mothers because it does not include discussion in which the child was not present. The mean amount of rehearsal for low-, moderate-, and high-severity groups was 0.95, 1.17, and 1.32, respectively, with the mean rehearsal rating across all participants at 1.15 ($SD = 0.40$). The low- and moderate-severity groups differed significantly from each other in amount of rehearsal, $t(85) = 2.04$, $p < .05$ ($ES = .04$). There was no significant difference between the moderate- and high-severity groups, $t(85) = 1.50$, *ns* ($ES = .01$). There was a significant positive correlation between amount of rehearsal and storm severity rating ($r = .30$, $p < .01$). Mothers who experienced a more severe hurricane tended to talk about the hurricane more. However, when the composite rehearsal score was used as a covariate in an analysis, with storm severity and type of recall as main factors, rehearsal was not a significant predictor of amount recalled, $F(1, 84) = 0.06$, $MSE = 0.11$, *ns* ($ES = 0$). As the sample size yielded sufficient power (82%) to detect a medium effect size (population $r = .30$) between rehearsal and propositions recalled, these results are not consistent with a meaningful relationship between rehearsal and recall. Further, the quadratic effect of storm severity, $F(1, 84) = 5.75$, $MSE = 0.11$, $p < .05$ ($ES = .05$), and the main effect of type of recall, $F(1, 85) = 32.03$, $MSE = 0.05$, $p < .01$ ($ES = .27$), were still significant after adjusting the mean propositions recalled for rehearsal. Thus, amount recalled as a function of stress and type of recall cannot be explained by rehearsal, as measured in this study. In fact, one would have predicted that the greater the rehearsal, the higher the recall. Hence, it is somewhat surprising that although amount of rehearsal was greater for mothers experiencing a high level of stress, this did not translate into higher levels of recall.

Rehearsal was also examined in more detail for the three time periods separately to determine whether greater rehearsal could account for the superior memory for T3 relative to T1 and T2. The mean amounts of rehearsal for T1, T2, and T3 were 0.67, 1.29, and 1.44, respectively. There was significantly more rehearsal in T3 than T2, $t(84) = 3.97$, $p < .01$ ($ES = .15$), and also significantly more rehearsal in T2 than T1, $t(84) = 10.88$, $p < .01$ ($ES = .58$). Rehearsal was not a significant predictor of differential recall across the three time periods, $F(1, 167) = 1.09$, $MSE = 0.02$, *ns* ($ES = 0$). As the sample size yielded sufficient power (81%) to detect a medium effect size (population $r = .30$) between rehearsal and propositions recalled for each of the three time periods, these results fail to demonstrate a meaningful relationship between rehearsal and number of propositions recalled at each time period. The main effect of time period was still clearly evident after the means were adjusted for effects of rehearsal, $F(2, 167) = 91.90$, $MSE = 0.02$, $p < .01$ ($ES = .52$).

For children, rehearsal was a significant predictor of recall across time periods, although it also did not impact the main effect of time period (see Bahrick et al., 1998). Although we did not find that rehearsal predicted the amount recalled across time periods for the mothers, we cannot confidently reject the possibility that rehearsal may have played a role. First, the rehearsal measure was specifically developed for the children and was not as accurate a measure for the mothers. Second, many investigators (e.g., Bohannon, 1988; Christianson & Loftus, 1991; Larsen, 1992; Neisser et al., 1996; Rubin & Kozin, 1984) have observed that self-reported frequency of rehearsal often does not correlate with memory performance in studies of emotional events. Neisser et al. (1996) claimed that asking their participants how much total time they spent "talking about the earthquake" was not a valid measure of rehearsal, as it did not say anything about their own personal experiences and was unrelated to accuracy of recall. Nevertheless, they claimed that each person had his or her own earthquake story and that these socially motivated narratives were most likely the basis of the high levels of accuracy documented in their study. Our mothers may also have had a hurricane story that adults would be more likely to repeat than children. Because we interviewed participants 2–5 months after the hurricane, these would have been solidly established narratives. In fact, this rehearsal may have compensated for what would otherwise be a drop in recall from moderate to high severity. Clearly, more precise methods of determining amount of rehearsal need to be developed to evaluate its role in memory of emotional events.

Retention interval. We were interested in how retention interval was related to the amount recalled. As mothers were interviewed between 60 and 201 days ($M = 113.96$, $SD = 28.73$) following the hurricane, we did not expect to observe the typical negative correlation between retention interval and amount recalled found for short delays. There was a nonsignificant correlation ($r = –.10$, *ns*) between retention interval and the log of the total propositions recalled, revealing no linear relationship between retention interval and amount recalled.

We were also interested in whether retention interval would predict the quadratic relationship between stress and recall. When retention interval was entered as a covariate, it was not a significant predictor of the amount recalled, $F(1, 92) = 0.07$, $MSE = 0.05$, *ns* ($ES = 0$). As the sample size yielded sufficient power (85%) to detect a medium effect size (population $r = .30$) between retention interval and propositions recalled, these results are inconsistent with a meaningful relationship between retention interval and recall. Further, the quadratic effect of storm severity, $F(1, 92) = 7.58$, $MSE = 0.05$, $p < .01$ ($ES = .07$), was still significant when the mean propositions recalled were adjusted for retention interval. The above analyses indicate that retention interval did not qualify the observed differences in recall, nor was retention interval related to recall in a linear way, most likely because all intervals were long, occurring after the initial forgetting leveled off.

In summary, the results of the analyses of covariance indicate that neither rehearsal nor retention interval significantly predicted the amount recalled by mothers about the hurricane-related events.

Conclusions

The present study assessed adult memory for Hurricane Andrew as a function of stress and compared results with those of our prior study (Bahrick et al., 1998) of children's memory for the same event. These studies addressed a number of limitations that have typically characterized previous research on memory and stress. We assessed recall of a natural disaster, a complex event that was extended in time. Previous studies of memory and stress have either been laboratory simulations that may not induce a high degree of stress, or they have been field studies that have looked

chiefly at memory for events surrounding the stressful event (e.g., FBM). Few studies of natural, extended events have examined recall of the stressful event itself (e.g., Pezdek, 2003a). Further, to date, there have been no other studies that carefully calibrated more than two levels of stress from the same natural disaster and examined their effects on recall. This is important because identifying at least three levels permits an examination of quadratic as well as linear relations between stress and memory. The present study differs from other investigations by objectively defining three levels of stress (based on amount of damage to the home) generated from the same event. This difference also provided a solution to the problem of defining a low-stress control group with content that was comparable with that of the high-stress group. Tests of the same event could be given to mothers who experienced a low, moderate, and severe storm, providing a basis for more comparable high- and low-stress comparisons.

Our parallel study of children's recall (Bahrick et al., 1998), conducted concurrently with the present study, demonstrated that preschoolers' recall of Hurricane Andrew showed an inverted-*U*-shaped curve, with the highest recall at moderate levels of stress. The present study assessed whether their mothers showed a similar quadratic function relating stress and recall. We also examined memory for different parts of the event (hurricane preparations, the hurricane itself, and its aftermath), for different types of content (descriptions, actions, and internal states), centrality of content, and type of memory test (free recall vs. prompted) as a function of the stress experienced, and we compared these findings with those of the children.

Our results demonstrate that mothers' overall recall of the hurricane showed a quadratic relation with stress. Recall increased from low to moderate levels of storm severity and then remained stable from moderate to high storm severity. This effect was observed even after recall was adjusted for the covariates of rehearsal and retention interval. Similarly, the recall of their children showed a quadratic relation with stress and demonstrated a significant rise from low to moderate levels of storm severity. However, the children showed a significant drop from moderate to high levels of storm severity.

Mothers also showed quadratic patterns relating stress and recall for different parts of the event: memories of the hurricane itself and memories of the hurricane aftermath. However, as predicted, memories of the hurricane preparations remained constant and did not differ as a function of stress. During the preparation period, mothers of all stress levels would have experienced the event as equally stressful given that it was not known where the storm would make landfall.

The quadratic relation between recall and stress and the failure to observe a significant drop in recall from moderate to high storm severity were evaluated in the context of inherent differences in the extent and amount of recallable material for individuals who experienced moderate versus high storm severity. Those who experienced high storm severity endured a hurricane aftermath that was significantly longer and had a significantly greater number of days without electricity, phone service, normal cooking, or water that was safe to drink (see Bahrick et al., 1998), as well as having a longer storm with more hurricane-related events during the storm itself. Further, there was evidence of more rehearsal at the high level of storm severity. These natural confounds would lead one to expect that memory for individuals who experienced high storm severity would be greater than for those who experienced low or moderate storm severity, given the greater amount of potentially recallable material in both the hurricane and aftermath periods. However, neither this greater amount of potential information nor the greater amount of rehearsal at high levels of stress translated into higher recall for the mothers in the high storm severity condition. Rather, these factors may have compensated for what would otherwise be a drop in recall at the high-stress level. This suggests that mothers, like their children, may experience some impairment in recall as a result of high stress. Future studies should address these confounds and when possible assess these differences and statistically control for amount of recallable material across stress levels.

For some content, mothers who experienced high stress showed a drop in recall. Mothers' recall of internal states reflected the curvilinear relationship with storm severity evidenced by their children for all content categories, showing a significant rise in recall from low to moderate storm severity and a significant drop in recall from moderate to high storm severity. Thus, for mothers, it appears that recall of emotional content was more impaired by high stress than other content. It should be pointed out that we cannot determine whether high levels of stress result in impaired memory or less willingness to report memories. Both possibilities could influence recall for emotional information at high stress levels. Mothers' recall of different types of information also mirrored that of their children: All participants recalled more actions than descriptions, and more descriptions than internal states.

We also examined memory and stress as a function of centrality of information. Researchers (e.g., Christianson, 1984; Heuer & Reisberg, 1992) have suggested that recall is higher for central information than peripheral information under stressful conditions. In the present study, and in our child study (Bahrick et al., 1998), the amount of recall was lower for PDR content (central information) than other information, and recall of PDR content and other content both showed quadratic functions with storm severity. Although we found consistency across age, the diverse definitions of centrality across research laboratories may well yield differing results.

Mothers and their children reported a great deal of information about the storm (*M*s = 520 and 154 propositions, respectively). Further, the amount of recall was significantly greater for prompted than free recall. In our interview procedure, we asked for free recall of the storm followed by specific prompts. Mothers and children reported a large amount of additional information following the specific prompts (*M*s = 310 and 110 propositions, respectively), even though participants of both ages reported having no more additional information to recall. This was true even for the most stressful portion of the event, the hurricane itself. Given that prompting elicited a large amount of additional information, we recommend that nonsuggestive prompts be used to maximize recall at all levels of stress and for individuals of all ages. Although the results of this study do not address the issue of the accuracy of this recalled information, evaluations of the children's transcripts by their mothers in Bahrick et al. (1998) suggest that the children's recall was accurate and contained few errors.

Overall, there are striking parallels between the recall patterns of adults and children. Both the child and adult data show greater recall of prompted than spontaneous information, peripheral than central information, actions than descriptions, and descriptions

than internal states. Further, they both showed an overall increase in recall from the low to moderate stress levels, with recall of the high-stress group remaining comparable with that of the moderate-stress group or dropping. It appears that high stress may negatively impact recall for both age groups.

Reisberg and Heuer (in press) pointed out that most trauma memories have considerable autobiographical significance, often involve familiar places and things, and are rehearsed more often in ways that aid memory. Because of this self-referenced framework and the fact that there is more potential information and evidence of more rehearsal at high stress levels, comparability of recall across moderate and high stress levels should not be taken as evidence that recall was not impaired by high stress. Rather, these factors may have simply offset a downward trend, caused by the stress inherent in the trauma.

Results of this research are relevant to the understanding of adults' and children's retrospective reports of stressful, temporally extended, natural events. Thus, they do not necessarily generalize to forensic situations involving brief unexpected events, such as an encounter with a stranger. Although adults recall a great deal more information about the hurricane than children, the results reveal parallel patterns between adults and children in terms of the relative amount of recall as a function of stress and for factors such as spontaneous versus prompted recall, content of recall, and centrality of information. These findings can inform forensic interviewers regarding the type and amount of recall expected from both children and adults in stressful circumstances involving familiar places and people and aid in choosing the most effective approaches for eliciting specific types of information.

References

Ackil, J. K., Van Abbema, D. L., & Bauer, P. J. (2003). After the storm: Enduring differences in mother–child recollections of traumatic and nontraumatic events. *Journal of Experimental Child Psychology, 84,* 286–309.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Bahrick, L. E., Parker, J. F., Fivush, R., Levitt, M. (1998). The effects of stress on young children's memory for a natural disaster. *Journal of Experimental Psychology: Applied, 4,* 308–331.

Bohannon, J. N. (1988). Flashbulb memories for the Space Shuttle disaster: A tale of two theories. *Cognition, 29,* 179–196.

Brown, J. M. (2003). Eyewitness memory for arousing events: Putting things into context. *Applied Cognitive Psychology, 17,* 93–106.

Brown, R., & Kulick, J. (1977). Flashbulb memories. *Cognition, 5,* 73–79.

Burke, A., Heuer, F., & Reisberg, D. (1992). Remembering emotional events. *Memory and Cognition, 20,* 277–290.

Byrne, C. A., Hyman, I. E., & Scott, K. L. (2001). Comparisons of memories for traumatic events and other experiences. *Applied Cognitive Psychology, 15,* S119–S133.

Christianson, S.-A. (1984). The relationship between induced emotional arousal and amnesia. *Scandinavian Journal of Psychology, 25,* 147–160.

Christianson, S.-A. (1989). Flashbulb memories: Special but not so special. *Memory and Cognition, 17,* 435–443.

Christianson, S.-A. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin, 112,* 284–309.

Christianson, S.-A., & Hubinette, B. (1993). Hands up! A study of witnesses' emotional reactions and memories associated with bank robberies. *Applied Cognitive Psychology, 7,* 365–379.

Christianson, S.-A., & Loftus, E. (1987). Memory for traumatic events. *Applied Cognitive Psychology, 1,* 225–239.

Christianson, S.-A., & Loftus, E. (1991). Remembering emotional events: The fate of detailed information. *Cognition and Emotion, 5,* 81–108.

Christianson, S.-A., & Nilsson, L. G. (1989). Hysterical amnesia: A case of aversively motivated isolation in memory. In T. Archer & L.-G. Nilsson (Eds.), *Aversion, avoidance, and anxiety: Perspectives on aversively motivated behavior* (pp. 289–310). Hillsdale, NJ: Erlbaum.

Clifford, B., & Hollin, C. (1981). Effects of the type of incident and number of perpetrators on eyewitness memory. *Journal of Applied Psychology, 66,* 364–370.

Clifford, B., & Scott, J. (1978). Individual and situational factors in eyewitness memory. *Journal of Applied Psychology, 63,* 352–359.

Cohen, J. (1988). *Statistical power analyses for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Conway, M. A. (1995). *Flashbulb memories.* Hillsdale, NJ: Erlbaum.

Deffenbacher, K. A. (1983). The influence of arousal on reliability of testimony. In S. M. A. Lloyd-Bostock & B. R. Clifford (Eds.), *Evaluating witness evidence* (pp. 235–251). New York: Wiley.

Deffenbacher, K. A. (1994). Effects of arousal on everyday memory. *Human Performance, 7,* 141–161.

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior, 28,* 687–706.

Easterbrook, J. A. (1959). The effect of emotion on the utilization and organization of behavior. *Psychological Review, 66,* 183–201.

Goodman, J., & Loftus, E. L. (1989). Implications of facial memory research for investigative and administrative criminal procedures. In A. Young & H. D. Ellis (Eds.), *Handbook of information on facial processing* (pp. 571–579). Amsterdam: Elsevier.

Heuer, F., & Reisberg, D. (1992). Emotion, arousal and memory for detail. In S.-A. Christianson (Ed.), *The handbook of emotion and memory: Research and theory* (pp. 151–180). Hillsdale, NJ: Erlbaum.

Jaccard, J. (1998). *Interaction effects in factorial analysis of variance.* Newbury Park, CA: Sage.

Jaccard, J. (2002). Analysis of variance frameworks in clinical child and adolescent psychology: Advanced issues and recommendations. *Journal of Clinical Child Psychology, 31,* 278–294.

Johnson, M., Foley, M. A., Suengas, A. G., & Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General, 117,* 371–376.

Kebeck, G., & Lohaus, A. (1986). Effect of emotional arousal on free recall of complex material. *Perceptual and Motor Skills, 63,* 461–462.

Larsen, S. F. (1992). Potential flashbulbs: Memories of ordinary news as a baseline. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: The problem of "flashbulb" memories* (pp. 32–64). Cambridge, England: Cambridge University Press.

Leippe, M. R., Wells, G., & Ostrom, T. M. (1978). Crime seriousness as determinant of accuracy in eyewitness identification. *Journal of Applied Psychology, 63,* 345–351.

Loftus, E. F., & Burns, T. E. (1982). Mental shock can produce retrograde amnesia. *Memory and Cognition, 10,* 318–323.

Neisser, U. (2003). New directions for flashbulb memories: Comments on the ACP issue. *Applied Cognitive Psychology, 17,* 1149–1155.

Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about the Challenger. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: The problem of "flashbulb" memories* (pp. 9–31). Cambridge, England: 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 versus hearing the news. *Memory, 4,* 337–357.

Pennebaker, J. W., & Harber, K. D. (1993). A social stage model of collective coping: The Loma Prieta earthquake and the Persian Gulf War. *Journal of Social Issues, 49,* 125–145.

Pezdek, K. (2003a). Event memory and autobiographical memory for the events of September 11, 2001. *Applied Cognitive Psychology, 17,* 1033–1045.

Pezdek, K. (Ed.). (2003b). Memory and cognition for the events of September 11, 2001 [Special issue]. *Applied Cognitive Psychology, 17*(9).

Pynoos, R. S., Frederick, C., Nader, K., Arroyo, W., Eth, S., Numez, W., et al. (1987). Life threat and posttraumatic stress in school age children. *Archives of General Psychiatry, 44,* 1057–1063.

Reisberg, D., & Heuer, F. (in press). The influence of emotion on memory in forensic settings. In M. P. Toglia, J. D. Read, D. F. Ross, & R. C. L. Lindsay (Eds.), *Handbook of eyewitness psychology: Vol. 1. Memory for events.* Mahwah, NJ: Erlbaum.

Rubin, D. C., & Kozin, M. (1984). Vivid memories. *Cognition, 16,* 81–95.

Safer, M. A., Christianson, S.-A., Autry, M. W., & Osterlund, K. (1998). Tunnel memory for traumatic events. *Applied Cognitive Psychology, 12,* 99–117.

Shore, J. H., Tatum, E. L., & Vollmer, W. M. (1986). Psychiatric reactions to disaster: The Mount St. Helen's experience. *American Journal of Psychiatry, 143,* 590–595.

Smith, M. C., Bibi, U., & Sheard, D. E. (2003). Evidence for the differential impact of time and emotion on personal and event memories for September 11, 2001. *Applied Cognitive Psychology, 17,* 1047–1055.

Steblay, N. M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior, 10,* 413–424.

Waganaar, W. A., & Groeneweg, J. (1990). The memory of concentration camp survivors. *Applied Cognitive Psychology, 4,* 77–87.

Wood, J. M., Bootzin, R. R., Rosenhan, D., Nolen-Hoeksema, S., & Jourden, F. (1992). Effects of the 1989 San Francisco earthquake on frequency and content of nightmares. *Journal of Abnormal Psychology, 101,* 219–224.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology of Psychology, 18,* 459–482.

Yuille, J. C., & Cutshall, J. L. (1986). A case study of eyewitness memory of a crime. *Journal of Applied Psychology, 71,* 291–301.

Yuille, J. C., & Cutshall, J. L. (1989). Analysis of the statements of victims, witnesses and suspects. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 175–191). Norwell, MA: Kluwer Academic.

Yuille, J. C., & Tollestrup, P. A. (1992). A model of the diverse effects of emotion on eyewitness memory. In S.-A. Christianson (Ed.), *The handbook of emotion and memory: Research and theory* (pp. 201–215). Hillsdale, NJ: Erlbaum.

Received September 8, 2005
Revision received June 6, 2006
Accepted June 6, 2006 ■