The Effects of Stress on Young Children’s Memory for a Natural Disaster

Lorraine E. Bahrick
and Janat Fraser Parker
Florida International University

Mary Levitt
Florida International University

The effects of stress on children’s long-term memory for a major hurricane were studied. Stress was objectively defined as low, moderate, or high according to the severity of damage to the child’s home. One hundred 3- and 4-year-old children received a structured interview 2–6 months following the hurricane. Older children recalled and elaborated more than younger children. Prompted recall was greater than spontaneous recall. There was a quadratic function, consistent with an inverted U-shaped curve, relating storm severity with overall as well as spontaneous recall. These findings can be applied to the effects of stress on the amount recalled by children giving retrospective accounts of temporally extended, naturalistic events.

On August 24th, 1992, Hurricane Andrew, a major Class IV hurricane with 175-mph winds, devastated much of the densely populated area of Dade County, Florida. The storm caused over 20 billion dollars of property damage and affected the lives of families and young children for months afterward. The present investigation assessed memory for events surrounding Hurricane Andrew in 3- and 4-year-old children. This event provided a unique forum for addressing many of the questions and issues raised by current developmental research on childhood event memory and its relation to stress in a naturalistic, personally relevant, and highly emotional event context.

Recent investigations of childhood memory have revealed that even quite young children are able to recall an impressive amount of accurate information about events they experienced (see Fivush, 1993 for a review). Further, memories are retained over very long time periods. For example, Hamond and Fivush (1991) interviewed 4- to 6-year-old children about a trip to Disneyworld that occurred either 6 or 18 months earlier. Children recalled a great deal about the event, approximately 40 propositions, and mothers judged virtually all the information to be accurate. Older children provided more elaborated and detailed verbal accounts than younger children, but amount of propositional information was not different. Other studies have also found accurate, long-term verbal recall of single, infrequent events such as the birth of a sibling.
MEMORY FOR A NATURAL DISASTER 309

(Sheingold & Tenney, 1982), a preschool evacuation due to fire (Pillemer, Picariello, & Fruet, 1995), trips to a zoo, circus, or birthday party (Fivush, Gray, & Fromhoff, 1987; Todd & Perlmutter, 1980), or a trip to an archaeological museum (Hudson & Fivush, 1991). In these studies, recall was evident across periods ranging from several months to several years, and in one study, across decades.

A growing body of research also indicates that older children recall more information than younger children in response to open-ended questions but not necessarily to more structured questions. Younger children rely more on specific questions, prompts, and structure to facilitate recall and can sometimes recall as much as somewhat older children when given sufficient prompts (Hamond & Fivush, 1991; Todd & Perlmutter, 1980; see Pillemer & White, 1989 for a review). Although many studies indicate even 1- to 3-year-olds show remarkably accurate and enduring memories (Bauer & Fivush, 1992; Bauer & Mandler, 1989; Bauer & Travis, 1993; Fivush, Kuebli, & Clubb, 1992; Hudson & Nelson, 1983; Price & Goodman, 1990; Slackman, Hudson, & Fivush, 1986), younger children also produce more errors of intrusion than older children (Howe, Courage, & Peterson, 1995) and can be more easily misled about the details of an event they witness (Ceci & Bruck, 1993; Ceci, Toglia, & Ross, 1987). This is of obvious practical importance when evaluating the potential for children to provide accurate testimony in court. Establishing criteria for evidence of memory may, however, differ depending on the purpose of the investigation (whether it is designed for generating developmental principles about how the memory system works, such as the present investigation, versus determining the reliability and accuracy of an eyewitness where accuracy of small details of an event may be crucial).

A critical concern from a forensic perspective has been the difficulty of comparing stress levels and memory tasks across studies. A variety of stress measures have been used and it is difficult to determine if high stress in one study is equivalent to high, moderate, or low stress in another study. This is especially true when subjective measures of stress are used. A further problem relates to differences in the type of memory tested. Memory has been tested in a variety of ways for many aspects of the events, ranging from free recall to identifying individuals in lineups and for aspects including central versus peripheral information and general knowledge versus specific details. Finally, it has been difficult to include a sufficiently wide range of stress such that a complete picture of the effects of stress on memory becomes apparent.

One of the most serious problems has been the difficulty of selecting an appropriate control group. Two general approaches have characterized the research described above. One approach has been to create a low stress control group by using an event that elicits little stress but is as comparable as possible to the stressful event in other ways (Goodman, Aman, & Hirschman, 1987; Goodman, Bottoms, Schwartz-Kenney, & Rudy, 1991; Goodman, Hirschman, Hepps, & Rudy, 1991) have argued that children who experience more arousal during medical procedures, such as inoculations, tend to show no decrement or enhanced memory for the event. In one study, their memory was superior to the control children who received a washable tattoo (Goodman, Hirschman, et al., 1991, Study 2). Other studies (e.g., Howe et al., 1995) found no relation between stress (as rated by parents) and amount recalled about an emergency room procedure either 3–5 days or 6 months later (see also Baker-Ward, Gordon, Ornstein, Larus, & Clubb, 1993; Howe et al., 1994). In contrast, Peters (1987, 1991) found that arousal either has no effect or hinders event memory. Children who were more highly stressed during a dentist visit, an inoculation, a fire alarm, or a visit by a stranger made a greater number of identification errors in lineups than those who were less stressed in the control condition, and stress was corroborated by measures such as pulse, blood pressure, and observational ratings. Others have also documented a negative relation between stress and memory under some conditions (Merritt, Ornstein, & Spicker, 1994; Vandermaas, Hess, & Baker-Ward, 1993).

One problem inherent in conducting this research has been the difficulty of comparing stress levels and memory tasks across studies. A variety of stress measures have been used and it is difficult to determine if high stress in one study is equivalent to high, moderate, or low stress in another study. This is especially true when subjective measures of stress are used. A further problem relates to differences in the type of memory tested. Memory has been tested in a variety of ways for many aspects of the events, ranging from free recall to identifying individuals in lineups and for aspects including central versus peripheral information and general knowledge versus specific details. Finally, it has been difficult to include a sufficiently wide range of stress such that a complete picture of the effects of stress on memory becomes apparent.

One of the most serious problems has been the difficulty of selecting an appropriate control group. Two general approaches have characterized the research described above. One approach has been to create a low stress control group by using an event that elicits little stress but is as comparable as possible to the stressful event in other ways (Goodman, Hirschman, et al., 1991, Experiments 1 and 3; Peters, 1997). Alternatively, researchers have chosen a stressful event for which there is individual variability in the degree of stress experienced and selected the children who are least anxious as the low stress control group (e.g., Goodman, Hirschman, et al.,
Each approach has advantages and limitations. The individual difference approach has the limitation that typically, a fairly narrow range of stress is sampled (a doctor or emergency room visit is rarely a positive experience). This restricts the likelihood of finding an effect if there is one present. Further, participants who experience these events as very stressful versus those who experience them as only mildly stressful may differ a priori in other important ways that may, in turn, be correlated with memory performance. For example, they may have different temperaments, general anxiety levels, or coping strategies. These preexisting differences make any observed differences in memory as a function of stress difficult to interpret. On the other hand, this approach has the important advantage that the memory test and task are identical across stress levels, and thus no confounds due to type of task arise.

The disadvantage of the approach in which participants are exposed to a low stress event that is similar to the stressful event is that the two events themselves differ. This raises the problem of confounds due to their lack of comparability. For example, the amount of prior knowledge has an important influence on what and how much children recall (Bjorklund, 1985; Chi & Ceci, 1987). Thus, if children have a more well-developed script or general knowledge base about one event than another, this could enhance amount recalled by providing better organization for details or by mistaking the general knowledge for specific event memories (see Clubb, Nida, Merritt, & Ornstein, 1993). A host of other concerns arise from the use of different events and makes it difficult to create truly comparable memory tests and events that differ only in the stress they elicit. Factors that impact memory performance must be equated across conditions, such as amount of central versus peripheral information presented and assessed, number of enabling or causal relations among components of events, and complexity of the events themselves. If different questions are asked for the different events, then the degree of structure and prompting versus open-endedness of questions, as well as overall complexity of the questions, must also be equated (see also Ornstein, Merritt, & Baker-Ward, 1993). For example, it is not possible to determine whether receiving an inoculation versus a washable tattoo (Goodman, Hirschman, et al., 1991), a teeth cleaning versus a surgical procedure at the dentist (Vandermaas et al., 1993), or a fire alarm versus the sound of a loud radio (Peters, 1997) is comparable in complexity, degree of general knowledge, or number of causal relations. For these reasons, it is difficult to be sure that observed differences in memory performance are a function of stress level and are not due to a priori differences in the memorability of the different events or the difficulty of the memory tests. This approach, however, often allows sampling of a much wider range of stress than the former approach. Ultimately, converging evidence from a variety of studies assessing memory across different events and settings is likely to provide us with a picture of the memory–stress relationship that is less biased by these factors.

Research on the relation between emotionality and memory in adults has yielded similarly inconsistent findings. Many studies have documented that emotional events are remembered with greater vividness and detail than neutral events (Christianson & Loftus, 1990; Pillemer, 1984; Rubin & Kozin, 1984). In other studies, emotionality has undermined memory for detail (Christianson & Nilsson, 1984; Clifford & Hollin, 1981; Clifford & Scott, 1978; Loftus & Burns, 1982; see Heuer & Reisberg, 1992 and Christianson, 1992 for reviews). Easterbrook (1959) proposed that arousal leads to a narrowing of attention and thus emotion may lead to better memory for central information but poorer memory for peripheral information. Consistent with this hypothesis, several studies found that central details were better remembered in emotional events than corresponding details in neutral events (Christianson & Loftus, 1987, 1990, 1991; Kebeck & Lohaus, 1986). Exceptions to this pattern have also been noted, however (e.g., Heuer & Reisberg, 1990, 1992). Thus, it appears that we are still far from being able to propose a clear, integrative account of how affect influences the memory system.

Hurricane Andrew, a natural disaster of enormous scope, was an extremely stressful event that provided a natural and opportune forum for investigating the relationship between memory and stress. It provided a number of excellent alternatives to many of the problems outlined
above. Families were subjected to a range of experiences from severe trauma to mild stress and inconvenience. Many residents feared for their lives and were forced to move from one room to another to avoid flying glass and debris as the storm decimated much of their home and its contents. Most of these families engaged in an extensive clean up period with no basic services for weeks and were then later forced to relocate while their homes were rebuilt. They experienced serious trauma. Other families who lived further away from the eye of the storm experienced a less severe storm with moderate levels of trauma and a great deal of inconvenience. They also received property damage (typically to the outside of their home, patio, and roof, and there was also water damage inside) and engaged in an extensive cleanup without basic services for several weeks. Typically, they repaired their homes without moving out. Those at the fringes of the storm experienced little direct stress, virtually no property damage, and only the inconvenience of a short period without basic services.

The hurricane provided an opportunity for investigating aspects of the memory–stress relationship that differ in important ways from those found in most of the prior studies. It provided a solution to the problem of defining the low stress control group. The same event and the same memory test could be given to children who experienced different levels of stress. All participants prepared in the same way for the hurricane, but some experienced an extremely traumatic event whereas others experienced only a heavy rain storm. Second, this research has the advantage that stress could be objectively defined according to the severity of the storm in each child’s neighborhood. This objective measure could also be correlated with more typical subjective ratings of stress. Third, the hurricane created a wide range of stress, from low to very severe, thereby allowing us a much broader view of the memory–stress relationship than is usually possible. This is particularly important given the discrepant findings across studies with respect to the facilitating versus debilitating effects of stress on memory. Fourth, the differential stress because of the storm occurred without regard for social and economic status. The average home value for families of low, moderate, and severe stress did not differ (see Results and Discussion, Table 1). Thus, our approach combines many of the advantages of the two prior approaches to the relationship between memory and stress, without being subject to the major limitations associated with each.

Prior research on effects of natural disasters and trauma on children has focused primarily on the nature of the stress and understanding its long term effects on psychological and social adjustment. Traumatic events investigated have included witnessing a sniper attack (Pynoos et al., 1987), a kidnapping (Terr, 1987), the murder of a parent (Malmquist, 1986), a dam collapse (Green, Korol, Grace, Vary, Leonard, Gleser, & Smienson-Cohen, 1991), a severe flood (Earls, Smith, Reich, & Jung, 1988), and a bushfire (McFarlane, 1987). Research on the effects of natural disasters has highlighted the long lasting effects of these traumas (e.g., Earls et al., 1988; McFarlane, 1987). Lonigan, Shannon, Taylor, Finch, and Sallee (1994) reported posttraumatic stress disorder (PTSD) symptoms in 5,687 children 3 months after they were exposed to hurricane Hugo. Similarly, Vernberg, La Greca, Silverman, and Prinstein (1996) analyzed 568 elementary school children’s reactions to Hurricane Andrew. More than 55% showed moderate to very severe levels of PTSD symptoms 3 months later. In spite of the recent surge of interest in the relationship between stress and memory, none of the research to date on the effects of natural disasters has focused on children’s memories for these disasters. Thus, the present study provides a bridge between the questions regarding memory and stress typically assessed in eyewitness research (in which stressful events such as visits to the doctor or dentist are assessed) and the natural disaster research (in which a much more devastating event with more far reaching consequences creates a much wider range of stress).

In the present study, we assessed memory for hurricane-related events in 3- and 4-year-old children and their mothers, 2 to 6 months following the hurricane. We chose to focus on preschoolers because of the controversy in the literature over the memory abilities of children this young. Stress was defined as the degree of storm exposure in each child’s neighborhood. Storm exposure was coded as high, moderate, or low on the basis of the amount and type of damage to the home. Structured interviews were given to chil-
children focusing on memory for three time periods: the prehurricane preparations, the hurricane itself, and the posthurricane recovery period. Given the previous literature, we assumed that children would be able to recall the hurricane experience. More critical, however, was how the level of stress might influence the amount and type of information reported.

Method

Participants

One hundred children aged 3 (n = 39; M = 1,357.3, SD = 133.0 days) and 4 years (n = 61; M = 1,738.7, SD = 232.3 days) at the time of testing and their mothers were recruited for participation through local preschools. There were a total of 51 boys and 49 girls (21 boys and 18 girls aged 3 years and 30 boys and 31 girls aged 4 years). Only the data of the children are reported here. Four additional children were interviewed but their data were rejected because the tape recording was incomprehensible (n = 1), the equipment failed (n = 2), or the experimenter failed to follow the interview protocol (n = 1). Children were recruited from an area spanning a 30-mile distance in Dade County, Florida, to include low (n = 18), moderate (n = 42), and severe (n = 40) hurricane exposure (see Results and Discussion, Storm Severity Classification, for details). Children were primarily from middle-class families with incomes between $40,000 and $80,000, and the mothers and fathers had an average of 15.4 (SD = 2.4) and 15.4 (SD = 2.6) years of education, respectively. The ethnic background of the children in the sample was 69.4% White, 20.4% Hispanic, 3.1% African American, and 7.1% other. All families had weathered the storm in their own homes.

Child Interview

The interview (see Appendix) 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). The interviewer initially asked an open-ended question about the general storm experience (e.g., “Tell me everything you can remember about the hurricane?”) 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 of these parts 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 your mommy and daddy do inside and outside the house to get ready for the hurricane?”). This was followed by specific item prompts (e.g., “Did your mommy and daddy do anything to the windows? What?”). The interview took approximately 30 min. Children then provided a brief subjective stress rating. They were assisted in rating on a 4-point scale how scared versus how good they and their mother felt during each of the three periods of the storm.

Parent Questionnaire Information

Mothers completed eight questionnaires including two measures of child behavior and six designed to measure the mother’s stress and behavior. The child measures included a modified version of The Child Frederick Reaction Index Form C (Frederick, 1985), consisting of 20 items designed to assess PTSD symptomatology in children and the “Child Well-Being Survey,” consisting of a 35-item questionnaire that we developed to assess changes in ordinary behaviors (such as sleeping, eating, attention, and emotionality) as a result of a traumatic event. Six measures were also included to assess the mother’s level of stress and changes in her behavior as a function of the hurricane. Because none of these additional measures of stress for the mother or child significantly predicted recall (either linearly or quadratically), they are not discussed further.

In addition, the mother filled out a questionnaire assessing demographic information, information about the consequences of the hurricane
to her home and disruption of services, and rehearsal of hurricane-related events. 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, 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, and T3). Four separate questions were asked including, "How often did your family talk to your child ..." or "... in front of your child . . .," about the prehurricane preparations, the hurricane itself, and the hurricane aftermath? A composite score was obtained by averaging across all questions.

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. If two subjects were named, the clause was parsed into two propositions ("My brother/ and my mother screamed loud"). If two action verbs were given for one subject, the content was also parsed into two propositions ("My brother screamed loud/then ran downstairs").

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 ("My uncle, who lived in New York, stayed with us;" "The name of the hurricane was Andrew;" "Bryan Norcross told everyone what to do."). Propositions were deemed irrelevant under the following categories: off-topic talk ("I got a present yesterday."), conversational fillers ("You know what?"), false starts ("I only . . . we were hiding . . ."), questioning the experimenter ("How did you know that?"), uncodable information (cannot understand what the child is saying), refusals ("That's all I'm gonna tell you."), repetitions (the child repeats in gist a prior statement), confirmation and denials (e.g., "yes," "nope"), and statements corrected later (child corrects own statement in another conversational turn). Interrater reliability was established for this stage among three independent coders (one primary and two secondary) on 27 of the 100 transcripts. The transcripts were systematically selected across interviewers to represent approximately an equal number of 3- and 4-year-olds and children of low, moderate, and high storm exposure. Reliability coding was done on one quartile (randomly selected) of each transcript. After extensive training with the coding system, each coder independently coded the selected quartile for 27 transcripts. Agreements between the primary and each secondary coder were counted when the phrase was parsed identically and was consistent in whether it was identified as hurricane relevant or irrelevant. The percentage of agreement be-
tween Coder 2 and the primary coder was 89.3% and between Coder 3 and the primary coder was 90.6%. After establishing this level of reliability, the three coders each coded approximately one third of the transcripts independently. Next, all hurricane-relevant propositions were coded according to four subclassifications (major content categories, preparation–damage–repair, time period, and type of memory) as well as for elaborations.

Coding for Content

All hurricane-relevant propositions were coded for content according to three major categories: actions, descriptions, or internal states. Each proposition was also coded as to whether or not it was relevant to hurricane preparations, damage, or repair.

Three major content categories. There were three major content categories: actions, descriptions, and internal states. 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., “My mommy cooked things on the grill;” “It blew the tiles off my roof”). 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, and have typically characterized descriptions (“The pool was all dirty;” “A big boom was heard”). When physical motion was implied but no causal agent was specified, the proposition was coded as a description (“My house got blown away”). Propositions were coded as internal states if they expressed emotion or affect (scared, happy, laughing, crying, or glad) or if they mentioned the following internal states: think, hungry, tired, know, dream, lucky, have a boo boo, and want (“I wanted to fix it”). Quotes were also coded but occurred too infrequently and by only a few children and were therefore later dropped from all analyses. Interrater reliability was calculated between two trained coders for one quartile of each of 27 transcripts. Percentage of agreement averaged 95.4%.

Preparation, damage, and repair content. Each proposition was further categorized as to whether it described preparation (“We bought lots of food”), damage (“The trees were all broken”), or repair and cleanup (“We picked up the shingles”) or whether it was not related to these activities. Interrater reliability was established between the same two coders as before, and percentage of agreement was 96.9%.

Coding for Elaborations

The same coders as above also counted all elaborations. Elaborations were defined as any adjective, adverb, possessive (except pronouns), modifying phrase (“near the pool,” “by himself,” “in the hurricane”), and relative clause (this, that, those, then, there, here, and when). For example, the proposition “a loud noise sounded by the car” was coded as having two elaborations, and the proposition “Nine people of my family stayed in the downstairs bathroom” was coded as having three elaborations. If the same elaboration was repeated, it was counted only once (“We had a lot, a lot of tiles down”). A given proposition might contain no elaborations or as many as four or five. The number of elaborations per proposition was calculated for each child. Interrater reliability was established as above, and each coder counted the number of elaborations for each proposition. Percentage of agreement averaged 99%.

Coding for Time Period and Specificity

All propositions were also coded for temporal and structural information.

Time period. Each proposition was rated as to whether the content was relevant to Time Period 1 (T1), the hurricane preparation period (“We brought in all the plants”), Time Period 2 (T2), the hurricane itself (“We were very hot in the closet”), Time Period 3 (T3), the hurricane recovery period (“and there were no cartoons”), or whether it was unrelated to a given time period (T4, “The hurricane was called Andrew”). Often, the coder had to use context to judge the time period.

Spontaneous versus cued recall. The interview was structured to facilitate investigation of spontaneous versus prompted recall because it
was composed of open-ended questions followed by increasingly more specific prompts. Propositions were thus coded according to the specificity of the interviewer's question that elicited the proposition. Four levels of prompting were defined: (1) the general open-ended question was defined as the most general level, (2) the temporal open-ended questions were defined as the next most general level, (3) category prompts were defined as more specific, and (4) the most specific level was a specific item prompt. For the purpose of investigating spontaneous versus prompted recall, all propositions elicited by open-ended questions (Levels 1 and 2) were classified as spontaneous memory, and all propositions elicited by category and specific item prompts (Levels 3 and 4) were classified as prompted recall.

Interrater reliability for time period and level of prompting was calculated between two new raters on one quartile of 27 interviews as before. Percentage of agreement averaged 93.6% for time period and 99.6% for level of prompting.

Results and Discussion

The major analyses explored the way in which the amount recalled about the hurricane was influenced by stress and child age. Stress was defined as storm severity and classified as high, moderate, or low (see below). Further, the total number of propositions generated by children and the degree of elaboration provided were analyzed as a function of time period, type of recall (spontaneous vs. prompted), and type of content, as well as age and storm severity. Finally, the effects of factors including rehearsal, retention interval, family income, and parental education on the main findings were examined.

Storm Severity Classification

Participants were classified into one of three stress levels (high, moderate, or low storm severity) depending on the mother's response to a 7-point questionnaire regarding storm exposure. If the storm actually penetrated the perimeter of the home while the family was inside, a rating of high severity was given. These occupants experienced breaking glass, flying objects, and physical danger. Most were required to flee from one room to another to avoid injury. A rating of moderate severity was given if during the storm, a great deal of property damage occurred, but the perimeter of the home was left mostly intact during the storm itself. 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 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. It is important to note that although the ratings of stress are objective, they are also relative to the range of stress experienced by children 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 approximately equal numbers of children who fell into the moderate (n = 42) and high severity (n = 40) groups, with fewer in the low severity condition (n = 18). Table 1 presents the consequences of the storm in terms of damage to the home and contents and interruption of basic services for children of the three stress levels, as well as analyses of variance (ANOVAs) assessing group differences for each variable. It is important to notice that the average value of the home did not differ as a function of hurricane severity. However, property damage and days without services differed significantly across groups in the direction expected, corroborating our hurricane severity measure.

Accuracy

Although accuracy of free recall for such a subjectively experienced, temporally extended natural event cannot be objectively assessed as in some studies of memory and stress in which records are available (e.g., visits to the doctor or dentist; Goodman et al., 1987; Goodman, Hirschman, et al., 1991), we nevertheless asked 25% of the mothers to judge their child's accuracy from a transcript of the child's interview. Mothers were asked to mark any content the child reported that was not completely true and to explain fully. Only 9 of the 25 transcripts were returned (many
families were understandably more concerned with rebuilding their homes). Still, these 9 transcripts can give us a general idea of the difficulty of judging accuracy and the type of errors mothers noted. Overall, the vast majority of propositions were deemed accurate. The mean number of propositions containing errors was only .033 per child, with a range of .013 to .078.

We developed a rough classification system for describing the range and types of errors children typically made. Many of the items marked inaccurate by parents nevertheless reflected memories of events that actually occurred. They included three categories: (1) time distortion \((n = 3\) of the 9 children; e.g., confused events before with after the storm), (2) inaccurate detail or part of a memory \((n = 7\); e.g., “Water came in the front door window,” but it was another window), and (3) exaggeration \((n = 5\); e.g., “Every window broke”). For the purposes of our study, we are not particularly concerned with these types of errors because they reflect memories of actual events. Other items marked as inaccurate by mothers were categorized as (4) self-contradictions \((n = 2\) or (5) illogical or impossible events \((n = 3\); e.g., “The soup was chasing us”). These, too, may not be errors in memory as much as errors in expression or thought.

Finally, three categories reflected some degree of memory distortion or inaccuracy and are more relevant to our purpose of assessing true memories. They included (6) inaccurate facts \((n = 6\); e.g., “Daddy went into the guest room,” but he actually could not get in), (7) wishful thinking \((n = 3\); e.g., “I got a lot of Kool Aid”), and (8) complete fabrications \((n = 2\). However, for the inaccurate facts, it was often not clear whether the child was actually in a position to know the information requested or not. These errors were more often made in response to a specific question asked by the interviewer. Finally, complete fabrications were rare. One child described cooking marshmallows, and another fabricated events including his grandfather in the closet (who was not present) and “the skeletons and their bones broke.” These examples highlight the complexity and difficulty of meaningfully assessing the accuracy of free recall of autobiographical information in young children. Further, it is not clear that the mother should be the ultimate standard against

### Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Overall</th>
<th>Low</th>
<th>Moderate</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value of home ($)</td>
<td>147,487</td>
<td>96,159</td>
<td>159,011</td>
<td>39,111</td>
</tr>
<tr>
<td>Damage to home ($)</td>
<td>34,625</td>
<td>35,938</td>
<td>34,312</td>
<td>34,201</td>
</tr>
<tr>
<td>Days without electricity</td>
<td>16.5</td>
<td>16.3</td>
<td>14.3</td>
<td>13.9</td>
</tr>
<tr>
<td>Days without running</td>
<td>5.5</td>
<td>7.6</td>
<td>9.0</td>
<td>10.4</td>
</tr>
<tr>
<td>Days without water</td>
<td>20.8</td>
<td>14.7</td>
<td>87.1</td>
<td>87.1</td>
</tr>
<tr>
<td>Days without phone</td>
<td>15.7</td>
<td>15.2</td>
<td>13.5</td>
<td>18.8</td>
</tr>
</tbody>
</table>

- Mean (M) values are reported for each group.
- Standard deviations (SD) are provided for each category.
- N indicates the number of participants in each group.
which to judge the child's memory for such a subjective, temporally extended event.

Amount Recalled as a Function of Age and Stress

The number of hurricane-relevant propositions generated by children as a function of age and storm severity is shown in Table 2. They ranged from 21 to 554, with an overall (unweighted) mean of 153.6 ($SD = 93.8$). In general, children spoke a great deal about the hurricane related events. Because the distribution was skewed, log transforms were performed and the log (base 10) of the total number of propositions generated served as our primary dependent measure (see Table 2).

Is there a relation between stress and amount recalled and is it affected by age? An ANOVA was conducted with child age (3 or 4 years) and storm severity (high, moderate, and low) on the log of the total number of hurricane-relevant propositions recalled. Results indicated a main effect of child age with 4-year-olds recalling significantly more than 3-year-olds, $F(1, 94) = 10.32, p = .002$ (see means in Table 2). This difference occurs even though developmental differences in the language production of 3- and 4-year-olds were partly minimized by excluding elaborations from this measure (such as the "big, bad" storm) and by evaluating elaborations separately (see below). Of even greater interest, a significant main effect of hurricane severity was found, $F(2, 94) = 3.14, p = .048$. Trend analyses were performed to assess the nature of the relationship between stress and memory. Results indicated a significant quadratic trend, $F(1, 94) = 5.84, p = .018$, relating the amount recalled and stress (as shown in Figure 1) and no linear trend, $F(1, 94) = 1.46, p > .1$. The moderate severity group recalled the most about the hurricane, whereas those in the low and high severity groups recalled the least.

Additional analyses were conducted for further evaluation of the nature of the quadratic trend. Given that the points defining the curve were somewhat arbitrarily demarcated (they represent an ordinal rather than an interval scale), it was not necessarily expected that recall from adjacent points would differ from one another. Rather, we were more interested in the shape of the curve.
Figure 1. Mean log number of propositions recalled as a function of storm severity.

Did retention increase and then decrease as a function of stress, or did it increase and then remain stable? Thus, we conducted analyses to evaluate the significance of the linear increase between recall of the low and moderate groups and the significance of the decrease from the moderate to high storm severity groups. Results of linear trend tests (one-tailed) indicated that the increase in recall from low to moderate storm severity groups was significant, $F(1, 94) = 5.86$, $p = .009$, and the decrease in recall from the moderate to high storm severity groups was marginally significant, $F(1, 94) = 2.19$, $p = .071$.

Thus, our data appear to be most consistent with an inverted U-shaped function relating the amount recalled and stress in young children. Apparently, when a broad range of stress is sampled, recall first increases and then decreases as stress intensifies. Further analyses were undertaken to determine the impact of secondary variables including type of recall and time period on this function.

Amount Recalled as a Function of Time Period and Type of Recall

We conducted more detailed analyses to examine memory as a function of time period (T1 = prehurricane preparations, T2 = the hurricane itself, and T3 = posthurricane recovery period) and type of recall (spontaneous vs. prompted). Most, but not all, propositions could be classified into one of the three time periods. Occasionally, children supplied relevant information that was not specific to a time period such as facts, opinions, or attitudes, and these were classified as unrelated to a time period (T4). The mean number of propositions classified into one of the three time periods was 143.24, compared with 153.6 total propositions generated about hurricane-related events.

All propositions were also classified as either spontaneous or prompted, depending on whether they had been elicited by open-ended questions versus category or specific prompts. Propositions elicited by the four open-ended questions were coded as spontaneous recall. Those elicited by category prompts and by specific item prompts were coded as prompted recall (see Procedure section for details).

Did Amount Recalled Differ Across Time Periods and for Spontaneous Versus Prompted Memories?

A four-way ANOVA with child age and storm severity as between-subjects factors and time period and type of recall as within-subjects factors was conducted on the log of the total propositions. Results indicated significant main effects of child age and storm severity as before and highly significant main effects of time period, $F(2, 188) = 139.7, p < .001$ (with untransformed means of 22.20, 63.18, and 57.86 for T1, T2, and T3, respectively), and type of recall, $F(1, 94) = 425.52, p < .001$ (with untransformed means of 33.28 and 109.96 for spontaneous and prompted recall, respectively.) Thus, significantly more information was elicited by prompted than by open-ended questions. Only two interactions reached significance: time period and storm severity, $F(4, 188) = 3.26, p = .014$, and time period and type of recall, $F(2, 188) = 14.68, p < .001$, and they are examined further below. Because
child age did not interact with any variables, further discussions will collapse across child age.

In What Way Did Recall for Events Surrounding Each of the Three Time Periods Differ?

We conducted post hoc tests comparing the log of the propositions recalled across the three time periods. Results indicated that children recalled significantly more propositions for T2 and T3 than for T1, \((p < .05, \text{Tukey's honestly significant difference})\). In fact, children recalled more than 2.5 times as many propositions about the hurricane itself than the preparation period. These differences should be viewed in the context of the nature and the extent of the three temporal events. The preparation period was approximately 2 days long and quite similar across children of different stress levels, the hurricane itself was approximately 8 h long, and the aftermath ranged from a few days to many weeks. Although the hurricane itself was the shortest period, much more was recalled about this period than the preparation period. Figure 2 displays the mean log number of propositions recalled for each time period as a function of storm severity. To help interpret the interaction between time period and severity, results were broken down according to time period and trend analyses were performed. A significant quadratic function was found across severity levels for memories surrounding T3, \(F(1, 94) = 8.21, p = .005\). Results of linear trend analyses also indicated that both the increase in recall from low to moderate storm severity conditions and the decrease from moderate to high storm severity conditions were significant, \(F(1, 94) = 7.07, p = .0005\) and \(F(1, 94) = 4.07, p = .023\), respectively. A marginally significant quadratic function was also found for memories surrounding T1, \(F(1, 94) = 2.88, p = .09\). Results of linear trend analyses revealed that the increase in recall from the low to moderate storm severity conditions was not significant, however, the decrease from moderate to high storm severity was significant, \(F(1, 94) = 4.70, p = .016\). The quadratic function for memories surrounding T2, however, was not significant, \(F(1, 94) = 1.89, p > .1\). Memory for this time period was also more confounded by extraneous factors that differed as a function of storm severity than memories for the other two periods. For example, in the low and moderate storm severity conditions, children were more likely to be able to sleep through portions of the storm (44% and 24%, respectively), than in the high severity condition (5%), and in the high severity condition, there were more distinctive events during the storm itself (the family typically moved from one room to another to avoid flying objects). This resulted in more potentially recallable salient events and more time awake for children in the high severity category and may have offset the tendency for recall to decrease with high levels of stress. In contrast, during T1, children all experienced the same kind of activities regardless of the level of hurricane severity because at that time, it was not known where the storm would make landfall. During T3, children in the low, moderate, and high severity groups may also have had increasingly more hurricane-related activities and thus more potentially recallable material. For example, children who experienced high storm severity had the largest number of days without
electricity, running water, or telephone, creating unusual circumstances for a longer period of time. The inverted U-shaped functions, both overall and particularly for the posthurricane recovery period, are thus striking given the expectation that the high storm severity group had more potentially recallable material, yet recall was the most attenuated.

**How Do Age and Stress Affect Spontaneous Versus Prompted Memories?**

Figure 3 displays the mean log number of propositions for open-ended versus prompted questions as a function of storm severity. In a three-way ANOVA with age, severity category, and type of recall, results revealed main effects of child age, storm severity, and type of recall as before, as well as the significant quadratic trend relating amount recalled with storm severity. However, there was also a significant interaction between storm severity entered as a quadratic and type of recall, $F(1, 94) = 4.53, p = .036$ (see Figure 3). These effects led us to examine and characterize in more detail the nature of memory under each type of recall condition separately. It was of particular interest to determine under what type of interview conditions children's memories of an event were most clearly differentiated as a function of stress. Separate analyses were conducted examining the log of the number of propositions generated by open-ended versus prompted questions as a function of age and storm severity.

**Spontaneous Memory: What Is the Relation Between Stress and Spontaneous Recall?**

Analyses for the spontaneous memories taken separately indicated a robust quadratic trend relating the amount recalled with hurricane severity, $F(1, 94) = 7.91, p = .006$ (see Figure 3). Further analyses revealed a significant linear increase in recall from the low to moderate storm severity groups, $F(1, 94) = 9.07, p = .002$, and a marginally significant decrease from the moderate to the high storm severity groups, $F(1, 94) = 2.18, p = .072$. As before, there were also significant main effects of child age, $F(1, 94) = 6.53, p = .012$, and hurricane severity, $F(2, 94) = 4.61, p = .012$. When each time period was examined separately, analyses indicated a significant main effect of storm severity for T3, $F(1, 94) = 4.09, p = .02$, and a significant quadratic function relating spontaneous recall and storm severity, $F(1, 94) = 8.18, p = .005$. Results of the linear trend analyses again revealed a significant increase in recall from the low to the moderate storm severity groups and a significant decrease in recall from the moderate to the high storm severity groups, $F(1, 94) = 6.06, p = .008$, and $F(1, 94) = 5.94, p = .013$, respectively. The quadratic function also approached significance for spontaneous recall of T2 events, $F(1, 94) = 3.59, p = .06$. These effects replicate and magnify those found for the overall results in which children in the moderate severity condition recalled the most about the event as a whole and especially about the posthurricane recovery period. When children's memory was not prompted, the amount recalled about the storm and especially the recovery period also tended to be greatest as a result of moderate stress and attenuated as a result of high stress. Thus, the relationship between spontaneous recall and stress was
Prompted Memory: What Is the Relation Between Stress and Prompted Recall?

Results of analyses for prompted memories were not as robust as those for the spontaneous memories. Although the main effect of child age was still evident when the prompted memories were considered alone, $F(1, 94) = 10.60, p = .002$, the main effect of hurricane severity was no longer apparent, $F(2, 94) = 2.18, p > .1$. Further, the quadratic trend relating amount recalled with hurricane severity approached but no longer reached significance, $F(1, 94) = 3.61, p = .06$ (see Figure 3). When prompted recall was examined for each time period separately; there were no significant main effects of storm severity (all $ps > .05$), but there was one significant quadratic trend as a function of severity for T3, as before, $F(1, 94) = 5.07, p = .03$. Results of linear trend analyses revealed a significant linear increase from the low to moderate storm severity groups, $F(1, 94) = 4.75, p = .016$, but only a marginally significant decrease from the moderate to the high storm severity conditions, $F(1, 94) = 2.17, p = .07$. Thus, overall, the analyses from the prompted memories show a pattern that is similar to those of the spontaneous and total memories, but effects of storm severity are attenuated. The main effect of storm severity and the quadratic trend found for the total recall measure were no longer observed when prompted recall was considered alone, even though prompting elicited more than 2.5 times as much recall as open-ended questions. In contrast, these effects were most evident when spontaneous recall was examined separately. Apparently, once spontaneous accounts have been given by children, trying to elicit more memory by using category prompts or specific cues results in a great deal more recall about the event for all stress levels and may diminish differences in amount recalled among children who experienced different levels of stress. Thus, by allowing children to report their memory in a free-recall context, differences in amount recalled as a function of stress may be maximized. By prompting memory, recall will be enhanced, especially for children who experienced more severe stress.

Relations Among Recall, Retention Interval, Rehearsal, and Socioeconomic Status (SES)

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, parent’s education, and home value), rehearsal of events surrounding the hurricane, and retention interval.

Could SES Predict the Relation Between Stress and Recall?

Table 3 depicts the mean family income and number of years of education for the mother and father as a function of storm severity. None of these measures significantly predicted amount recalled when used as covariates, $F(1, 91) =$

Table 3
Socioeconomic Status: Means and Standard Deviations for Family Income and Number of Years of Education as a Function of Storm Severity

<table>
<thead>
<tr>
<th>Income and education</th>
<th>Low (n = 18)</th>
<th>Moderate (n = 41)</th>
<th>High (n = 39)</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Family income</td>
<td>5.19</td>
<td>1.44</td>
<td>5.10</td>
<td>2.03</td>
</tr>
<tr>
<td>Father's education</td>
<td>14.98</td>
<td>2.45</td>
<td>16.59*</td>
<td>2.73</td>
</tr>
<tr>
<td>Mother's education</td>
<td>15.33</td>
<td>2.57</td>
<td>15.90</td>
<td>2.05</td>
</tr>
</tbody>
</table>

Note. For income, 1 = under $10,000, 2 = $10,000-$19,999, 3 = $20,000-$29,999, 4 = $30,000-$39,000, 5 = $40,000-$49,000, 6 = $60,000-$79,000, and 7 = Over $80,000.
0.52, p > .1, F(1, 91) = 1.84, p > .1, and F(1, 90) = 0.68, p > .1, for income, mother's education, and father's education, respectively. Further, the main effects of the child's age, storm severity, and type of recall, as well as the quadratic functions relating storm severity and amount recalled were all significant even after the means were adjusted for the effects of each of these variables (all ps < .05). Further, the value of the family home (see Table 1) was also used as a covariate and did not predict recall, F(3, 72) = 0.58, p > .1.

**Could Rehearsal Predict the Quadratic Relation Between Stress and Recall?**

Might the quadratic function relating amount recalled and stress be mediated by differential rehearsal across the storm severity groups? Perhaps families who experienced moderate storm severity discussed the hurricane-related events more than families who experienced high or low storm severity. It is important to recall that our questionnaire assessed rehearsal on a 3-point scale ranging from 0 (none) to 1 (once a day) or 2 (several times a day) for a number of different questions and time periods, and a composite score was derived. The mean rehearsal rating across all participants was 1.5 (SD = 0.40). It was 1.33, 1.55, and 1.57 for the low, moderate, and high storm severity groups, respectively. There was a significant positive correlation between amount of rehearsal and storm severity rating (r = .39, p < .01). Families who experienced a more severe hurricane tended to talk about the hurricane more. When the composite rehearsal score was used as a covariate in an analysis with the child's age, storm severity, and type of recall as main factors, rehearsal was not a significant predictor of amount recalled, F(1, 84) = 0.04, p > .1. Further, all effects previously found for child age, storm severity entered as a quadratic, and type of recall were still significant even after adjusting the means for rehearsal, F(1, 84) = 8.8, p = .004, F(1, 84) = 7.2, p = .009, and F(1, 85) = 27.42, p < .001, respectively. Thus, amount recalled as a function of stress, age, and type of recall cannot be explained by differential rehearsal. 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 T2 and T3 relative to T1. The mean amounts of rehearsal for T1, T2, and T3 were 1.02, 1.75, and 1.68, respectively. Rehearsal was a significant predictor of differential recall across the three time periods, F(1, 167) = 4.36, p = .038, however, the main effect of time was still evident even after the means were adjusted for effects of rehearsal, F(2, 167) = 38.34, p = < .001.

**Could Retention Interval Predict the Quadratic Relation Between Stress and Recall?**

It is important to recall that children were interviewed between 60 and 179 days (M = 102.7, SD = 24.02) following the hurricane, and thus we did not expect to observe the typical negative correlation between retention interval and amount recalled found for short delays. When retention interval was entered as a covariate, it was not a significant predictor of the amount recalled, F(1, 93) = 0.65, p > .1. Further, all main effects of age, type of recall, and the quadratic function were still significant when the means were adjusted for retention interval (all ps < .05).

**How Was the Retention Interval Related to the Amount Recalled?**

A correlation between retention interval and the log of the total propositions recalled revealed a nonsignificant correlation (r = .15). Thus, there was no linear relation between retention interval and amount recalled. Further, when children were classified into three retention interval groups of approximately equal number, short (60–92 days), intermediate (93–108 days), and long (109–179 days), those with the shortest interval recalled the most (M = 2.21), but those with the longest interval recalled the next most (M = 2.13), and those with the intermediate interval recalled the least (M = 1.80). The analyses above indicate that the 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.

Taken together, the results of the analyses of covariance indicate that none of the covariates, including four measures of SES, rehearsal, or
retention interval, significantly predicted the amount recalled by children about the hurricane-related events.

Relation Between Stress and Elaborations

Are Memories of Stressful Events More or Less Elaborated Than Memories of Less Stressful Events?

The mean number of elaborations per proposition served as the dependent variable in these analyses. Overall, children generated 0.88 (SD = 0.26) elaborations per proposition, with an average of 0.81 (SD = 0.23) for 3-year-olds and 0.95 (SD = 0.29) for 4-year-olds. An ANOVA with child age and storm severity as main factors revealed a significant main effect of age, $F(1, 94) = 4.86, p = .03$, and no effect of storm severity or interaction ($ps > .1$). Four-year-olds produced significantly more elaborations than 3-year-olds. Further, when storm severity was entered as a quadratic, there was no significant effect on elaborations, $F(1, 94) = 1.52, p > .1$. In general, recall was not more elaborated for children with moderate storm exposure than those with low or high exposure.

Did Children Elaborate More When Describing Events That Occurred During One Time Period Over Another?

Results of a three-way ANOVA with time as a repeated measures factor revealed a significant main effect of time period, $F(2, 184) = 13.03, p < .001$, with means of .72, .85, and .96 for T1, T2, and T3, respectively. Apparently, children generated more elaborations when they were recalling events surrounding the hurricane itself and the aftermath as compared with the preparation period.

How Is Elaboration Related to Spontaneous Versus Prompted Recall?

A three-way ANOVA was conducted with type of recall as a factor along with child age and storm severity. Results indicated no significant main effect of type of recall, $F(1, 93) = 1.11, p > .1$, or interactions. However, when storm severity was entered as a quadratic, there was a marginally significant effect, $F(1, 93) = 3.16, p = .08$. This quadratic function was significant when spontaneous recall was examined separately, $F(1, 93) = 4.89, p = .029$, and was not significant for prompted recall ($p > .1$). Thus, when responding to open-ended questions, children tended to elaborate more under the moderate storm severity conditions and less under the high and low severity conditions.

Taken together, these analyses indicate that 4-year-olds produced more elaborations than 3-year-olds and that children elaborated most when they were describing the hurricane aftermath and least when they were describing the preparation period. Overall, the degree of elaboration was not systematically related to storm severity. However, under certain conditions, especially for spontaneous recall, there was a quadratic function relating the degree of elaboration with storm severity.

Memory Content as a Function of Age and Stress

What Kind of Information Is Recalled by Children and Does This Differ as a Function of Age and Stress?

Table 4 displays the mean number of propositions classified into each content category (actions, descriptions, and internal states). Children generated a mean of 51.92 descriptions, 86.16 actions, and 13.18 internal states. A three-way ANOVA with child age and storm severity as between-subjects factors and content category as a within-subjects factor was conducted on the log of the total hurricane-relevant propositions. Results indicated significant main effects of child age, $F(1, 94) = 10.97, p = .001$, and storm severity entered as a quadratic, $F(1, 94) = 5.51, p = .021$, as in previous analyses, and a significant main effect of content category, $F(2, 188) = 677.99, p < .0001$. Post hoc analyses indicated that children produced significantly more actions than descriptions ($p < .05$) than internal states ($p < .05$). However, there were no interactions between content category and child age or between content category and storm severity ($ps > .1$). Apparently, age and stress level had little influence on the relative amounts of actions,
Table 4
Memory Content: Mean Number of Propositions, Standard Deviations, and Log Transforms (Base 10) for the Means as a Function of Storm Severity and Content Category

<table>
<thead>
<tr>
<th>Content category</th>
<th>Low</th>
<th>Moderate</th>
<th>High</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>log</td>
<td>M</td>
</tr>
<tr>
<td>Descriptions</td>
<td>45.95</td>
<td>35.11</td>
<td>1.54</td>
<td>57.79</td>
</tr>
<tr>
<td>Actions</td>
<td>79.71</td>
<td>68.90</td>
<td>1.77</td>
<td>101.04</td>
</tr>
<tr>
<td>Internal states</td>
<td>11.43</td>
<td>9.91</td>
<td>0.96</td>
<td>14.97</td>
</tr>
</tbody>
</table>

Note. The marginal means shown here were calculated by weighting the cells equally to be consistent with the analyses of variance used.

descriptions, and internal states generated by children.

Did Memories About Hurricane Preparations, Damage, and Repair Vary as a Function of Stress?

The content of children's recall was classified as related or unrelated to hurricane preparation, damage, or repair and recovery (PDR). Table 5 displays the mean number of propositions recalled by 3- versus 4-year-olds about hurricane PDR as a function of storm severity. Did children who were more severely stressed talk less about the direct effects of the storm (PDR)? Was discussion of this central content related to stress by a quadratic function? A three-way ANOVA with age, storm severity, and content (PDR vs. other content) was conducted on the log of the total hurricane propositions recalled. In addition to the main effects of age and storm severity observed previously, children showed a main effect of PDR content, $F(1, 94) = 256.40, p < .001$. There was significantly more information recalled about material unrelated to PDR ($M = 115.3$) than material related to this content ($M = 38.1$). Of greater interest were the two interactions that emerged. One was between type of content and storm severity, $F(1, 94) = 7.13, p = .001$. It reflects the fact that although there was a significant quadratic function relating storm severity and recall for both PDR and non-PDR content, $F(1, 94) = 4.96, p = .028$, and $F(1, 94) = 5.24, p = .024$, respectively, there was also a strong linear component for recall of PDR content, $F(1, 95) = 11.74, p = .001$. There was no evidence of a linear component for non-PDR content. This effect is better understood by the

Table 5
Memory Content: Mean Number of Propositions, Standard Deviations, and Log Transforms (Base 10) for Mean Recall About Hurricane Preparation, Damage, or Repair (PDR) Versus Other Content as a Function of Age and Storm Severity

<table>
<thead>
<tr>
<th>Storm severity</th>
<th>Low (n = 18)</th>
<th>Moderate (n = 42)</th>
<th>High (n = 40)</th>
<th>Overall (N = 100)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>log</td>
<td>M</td>
</tr>
<tr>
<td>PDR Content</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 years (n = 39)</td>
<td>28.71</td>
<td>39.80</td>
<td>1.15</td>
<td>32.58</td>
</tr>
<tr>
<td>4 years (n = 61)</td>
<td>33.09</td>
<td>32.61</td>
<td>1.30</td>
<td>53.48</td>
</tr>
<tr>
<td>Overall (N = 100)</td>
<td>30.90</td>
<td>36.38</td>
<td>1.23</td>
<td>43.03</td>
</tr>
<tr>
<td>Other Content</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 years (n = 39)</td>
<td>75.29</td>
<td>52.56</td>
<td>1.78</td>
<td>116.90</td>
</tr>
<tr>
<td>4 years (n = 61)</td>
<td>141.55</td>
<td>102.65</td>
<td>2.06</td>
<td>151.48</td>
</tr>
<tr>
<td>Overall (N = 100)</td>
<td>108.42</td>
<td>81.56</td>
<td>1.92</td>
<td>134.19</td>
</tr>
</tbody>
</table>

Note. The log transforms may result in numerical differences with respect to the observed means. The marginal means shown here were calculated by weighting the cells equally to be consistent with the analyses of variance.
other interaction that emerged, a three-way interaction between type of content, storm severity, and child age, $F(1, 94) = 3.26, p = .04$. Three-year-olds showed an increasing linear function for recall of PDR content as storm severity increased, $F(1, 36) = 7.21, p = .011$, whereas 4-year-olds showed the typical quadratic function, $F(1, 58) = 5.66, p = .021$.

Thus, for memories related to PDR, only the 4-year-olds showed a quadratic function relating recall and stress. Three-year-olds showed an increasing tendency to talk about the PDR content as hurricane severity increased. Although these findings were not designed a priori to be relevant to the literature on attentional narrowing and stress (e.g., Easterbrook, 1959), they do suggest that there are age differences in the patterns of recall across stress levels for different types of memory content. If PDR content is conceptualized as exclusively central information, then the present results suggest some evidence of attentional narrowing with increasing stress for 3-year-olds but not 4-year-olds. That is, they recalled more about hurricane PDR as storm severity increased, and less about content unrelated to preparation, damage, and repair. It should be noted, however, that, the remaining (nonPDR) content cannot be considered peripheral because it included highly relevant information about other aspects of the storm such as the sounds it made, the people who were present, what they did and said, feelings about the storm, and so forth, as well as peripheral information. In the literature, definitions of central information vary widely, ranging from “weapon focus,” to material related to the critical event, to material that is centrally located in the visual field in photos, slides, and videotaped presentations. A more objective definition of central versus peripheral information is needed if meaningful comparisons are to be made across studies.

Conclusions

The results of this study present remarkable evidence of long, detailed accounts of a stressful experience that occurred 2 to 6 months earlier. Three and 4-year-old children recalled an average of 153 propositions about Hurricane Andrew. This is far greater than the amount of information reported in prior studies in which young children recounted personal events after lengthy delays (e.g., Fivush et al., 1987; Hammond & Fivush, 1991; Ornstein et al., 1995). This may be due in part to the duration of the hurricane event and the amount of discussion about it. The hurricane event was temporally extended, lasting weeks from preparation to posthurricane recovery, especially for the moderate and high stress groups, and it was a topic of discussion for weeks afterwards. In fact, families reported that they discussed the event in front of their child between once a day and several times a day across the 2- to 6-month retention interval. Thus, rehearsal may in part account for the impressive amount of recall. Still, these results provide clear evidence that very young children report a great deal of information about highly emotional and stressful experiences. Further, although we were unable to address the issue of accuracy objectively, the little data we were able to collect showed that mothers judged their children's reports to be highly accurate.

In this research, we were able to resolve a number of problems typically inherent in conducting naturalistic studies on memory and stress. Stress was objectively defined in terms of the degree of hurricane severity and children were classified into three stress levels on this basis. Thus, the problem of obtaining an appropriate low stress control group was resolved by including families who prepared for the hurricane but experienced only a typical storm. Further, because the hurricane provided an opportune and objective way of measuring stress, we did not have to rely on subjective ratings given by young children. In addition, our results are not affected by differences across stress levels in the memory test because a single test was given of a distinct event that varied in intensity. Finally, there were no systematic individual differences across stress groups because assignment to groups was not systematic, and prior knowledge of the event was minimal and not related to stress level.

In this context, our research uncovered a quadratic relation between stress and memory. Children who experienced a moderate level of stress recalled the most information whereas those who experienced high or low stress recalled the least. Overall, recall showed a significant linear increase from the low to moderate conditions and a marginally significant decrease from
the moderate to high storm severity conditions. When memory for the three parts of the storm (the hurricane preparation period, the storm itself, and the recovery period) was examined separately, we found that the quadratic function was significant and most pronounced for memories surrounding the recovery period, marginally significant for memories of the preparation period, and not significant for memories of the hurricane itself. This lack of significance for memories of the storm itself may have occurred because for this period, recall was more confounded by extraneous factors than for the other time periods. For example, potential decreases in memory from the moderate to the high storm severity groups may have been offset by the increase in the amount of hurricane-related activities during the storm itself for the high severity group (moving from room to room to avoid flying objects) and by the fact more children in the moderate and low severity groups were able to sleep through portions of the storm. This would result in more potentially recallable material during the storm itself for children in the high storm severity condition. Moreover, analyses indicated that for the two time periods where the quadratic function was observed, there was a significant linear decrease in recall from the moderate to the high storm severity conditions, and for memory of the recovery period, a significant linear increase in recall from the low to moderate storm severity conditions. Taken together, the findings of a quadratic function relating amount recalled and storm severity appear to be most consistent with an inverted U-shaped curve where memory first increases and then decreases as a function of stress.

When spontaneous recall was examined separately, the quadratic function was most pronounced. Children were most likely to report differing amounts of material as a function of stress when they answered open-ended questions. Here, too, the linear increase in recall from low to moderate storm severity conditions was significant and the linear decrease from moderate to high storm severity was marginally significant. Similar to results of overall recall, the quadratic function and the inverted U-shaped relation between recall and stress was most evident in the spontaneous memories for the posthurricane recovery period. The amount recalled showed both a significant linear increase and then a significant decrease from low to moderate to high storm severity groups. Thus, results of spontaneous recall parallel and magnify the effects found for overall recall. The quadratic function also approached significance for prompted recall but was not as robust.

The quadratic function was also observed under a variety of other conditions and for both 3- and 4-year-old children. Although the degree of memory elaboration was not systematically related to stress when all memories were considered together, when spontaneous memory was considered alone, the degree of elaboration was related to stress by a quadratic function. In addition, the amount recalled was related to stress by a quadratic function regardless of the content of the children's recollections. When recall of material related to the most stressful aspects of the event (PDR) was examined separately, the quadratic function was found. This function was also obtained for recall of material unrelated to PDR, and thus did not appear to depend on the centrality of the content. Finally, the curvilinear function was observed regardless of whether the content of children's memories included actions, descriptions, or internal states. Thus, the quadratic function relating memory and stress was robust and was observed for children of both ages under a variety of conditions. Further, the quadratic relation between recall and stress was not qualified by the family's SES (mother's or father's education, value of the home, or family income), the retention interval, or the amount of rehearsal. In fact, the rehearsal of storm-related events was greatest for the families who experienced the greatest storm severity. That recall was attenuated for the high stress group despite the greater level of rehearsal also attests to the robustness of the curvilinear relation between memory and stress.

For most measures, our research revealed that the patterns of recall were similar across age, and 4-year-olds recalled consistently more information than 3-year-olds. At both ages, children showed a curvilinear relation between amount of recall and stress, remembering the most under the moderate storm severity condition. At both ages, they also remembered more actions, than descriptions, than internal states, and they recalled more about the hurricane aftermath and the hurricane itself than they did about the hurricane prepara-
tion period. It is interesting that prompting memory with category or specific prompts led to more than a twofold increase in the number of propositions reported over spontaneous memory, and this was true at both ages as well. One of the few exceptions to the parallel pattern across age was the recall of information related to the more stressful aspects of the event, hurricane PDR. Overall, there was a significant quadratic function relating stress and amount recalled. However, when recall was broken down by age, the 4-year-olds showed a quadratic function and the 3-year-olds showed an increasing linear function. The data of the 3-year-olds were most consistent with an attentional narrowing hypothesis.

This research, thus, provides important and practical information about the nature of memory for highly stressful experiences. The stress generated by the hurricane was clearly more extreme than can be studied in experimentally controlled or laboratory studies of stress. The type of trauma generated by a hurricane is extended over time and diverse, including both physical and psychological stress. Families in the high severity condition feared physical injury, experienced unpredictability and helplessness during the storm, and experienced an extended period of physical and emotional hardship for several days prior to and several weeks after the storm. Thus, the results of this research can be applied to our understanding of children's retrospective reports of temporally extended, stressful events that occur outside the laboratory setting. They highlight the type and amount of material presented by children experiencing different levels of stress and how this changes with prompting versus open-ended questioning and as a function of centrality of information. These findings can educate forensic interviewers as to the most effective approach to use as a function of the child’s stress. Further, to better evaluate the trade-offs between open-ended versus prompted questioning as a function of stress, future research must assess the impact of these variables on the accuracy of the information reported.

In contrast with prior research on the relation between memory and stress, the range of stress experienced by children in our study is most likely more broad than that of most lab and field studies of event memory. During the hurricane, some children actually feared for their lives and personal safety whereas others experienced only a normal storm. Prior research has often classified children into only two levels of stress or assessed individual differences in stress to a single event, restricting the range of stress sampled. Some studies have found that stress does not impact memory or can even enhance memory (e.g., Goodman et al., 1987; Goodman, Hirschman, et al., 1991), whereas others have found that it impairs memory (e.g., Peters, 1987, 1991). By classifying children according to three levels of stress, we were able to identify a quadratic relation between memory and stress most consistent with an inverted U-shaped function that cannot be revealed by assessing memory across only two stress levels or by sampling a narrow range of stress.

In summary, children reported a great deal of information about a stressful event and the level of stress influenced the amount recalled several months later. However, to assert that moderate levels of stress enhance memory and high levels of stress disrupt memory would be an oversimplification. It is apparent that this highly emotional and important event elicited an impressive amount of recall, even by those in the high storm severity condition, relative to studies of other events recounted by children. Further, the fact that children under all stress levels recounted much more information after being prompted points out that additional memories of even the most stressful events can be elicited by prompting. Thus, the finding of a quadratic relation between memory and stress should be viewed in the context of the large number of memories reported by children about a highly emotional and temporally extended, naturalistic event.

References


(Appendix follows on next page)
Appendix

Hurricane Memory Interview

*Open-ended general question*

"Remember the really big hurricane we had here? I'm really interested in what kids like you remember about that hurricane. I've already talked to some kids about what happened and they told me all kinds of interesting things. I bet you remember a lot about the hurricane, too. Can you think really hard about the hurricane and tell me everything that you can remember about it?" (Follow up with nondirective prompts [NDP]; "What else?", "Anything more?", "Anything else?", "What?", and "Any other things?")

"You really remembered a lot about what happened. That's great!"

*Prehurricane preparations*

"Can you tell me what you and your family did to get ready before the hurricane came? Think about all the things that you and your mommy and daddy did to get ready for the hurricane. Tell me everything you can. What was the very first thing that your family did?" (NDP)

(If the child did not mention the following items, the specific questions were asked and were followed by NDPs.)

"What did your mommy and daddy do outside the house?"

"Did they bring anything inside? Plants? Lawn furniture?"

"What did your mommy and daddy do inside the house? What about the rugs? Furniture? Water?"

"What special things did your mommy and daddy buy to get ready for the hurricane?" (NDP)

"What about groceries?"

"Did you help your family get ready for the hurricane? What did you do?"

"Was anybody at your house before the storm besides who usually lives there? Did they help you get ready?"

"What was the most scary thing about getting ready for the hurricane?"

"What was the most fun thing about getting ready for the hurricane?"

*The hurricane itself*

"What do you remember about what happened when the hurricane finally came? Tell me everything you can." (NDP) "And then what?"

(If the child did not mention the following items, the specific questions were asked and were followed by NDPs.)

"Were you awake during any of the hurricane? What part? Tell me about it."

"Where did you stay in the house?"

"Where did your mommy and daddy stay in the house?"

"Did you hide or put something over you?"

"What was it like in the house?"


"What was it like outside during the hurricane?"

"Could you see anything?" Wind? Noises? What were the noises from?

"Was anybody at your house during the hurricane besides who usually lives there?"

"What was the scariest thing during the hurricane?"

"What was the most fun thing during the hurricane?"

*Posthurricane recovery period*

"Okay, now I want you to tell me all the things you can remember from after the hurricane was over. Tell me everything that you can think of."

(If the child did not mention the following items, the specific questions were asked and were followed by NDPs.)


"Now try to remember what it was like when you first looked outside and the hurricane was over. What did it look like? What was in your yard that was different?"

"Trees? Junk?"

"What did your mommy and daddy do?"

"Did you help with anything? What did you do?"

"Think about what happened those first few days after the hurricane. Tell me what that was like. Think of all the things that were different."
“Did you stay in your own house?” If no, “Where did you go?” If yes, “Did someone come to stay with you?”
“What was the scariest thing after the hurricane was over?”

“What was the most fun thing after the hurricane was over?”

Received October 28, 1996
Revision received January 23, 1998
Accepted January 30, 1998