CHILD DEVELOPMENT

Child Development, May/June 2014, Volume 85, Number 3, Pages 1292–1305

Stress, Interviewer Support, and Children's Eyewitness Identification Accuracy

Elizabeth B. Rush, Jodi A. Quas, and Ilona S. Yim
University of California, Irvine

Mariya Nikolayev
George Mason University

Steven E. Clark and Rakel P. Larson
University of California, Riverside

Few studies have investigated how stress affects eyewitness identification capabilities across development, and no studies have investigated whether retrieval context in conjunction with stress affects accuracy. In this study, one hundred fifty-nine 7- to 8- and 12- to 14-year-olds completed a high- or low-stress laboratory protocol during which they interacted with a confederate. Two weeks later, they attempted to identify the confederate in a photographic lineup. The lineup administrator behaved in either a supportive or a nonsupportive manner. Participants who experienced the high-stress event and were questioned by a supportive interviewer were most accurate in rejecting target-absent lineups. Results have implications for debates about effects of stress on eyewitness recall, how best to elicit accurate identifications in children, and developmental changes in episodic mnemonic processes.

During the past two decades, findings from an expansive body of research concerning adults' eyewitness identification abilities have not only had a tremendous impact on legal decisions and the pursuit of justice when witnesses are questioned about crimes, but also broadened understanding of the complex ways in which memory for faces is influenced by the context within which the faces are encoded and retrieved (see Clark, 2012; Cutler & Penrod, 1995; Wells et al., 1998). By comparison, very few studies have examined children's and adolescents' eyewitness identification capabilities (see Pozzulo & Lindsay, 1998). However, they also experience crime, at times even more often than adults (Bureau of Justice Statistics, 2008). Moreover, their memory abilities, including for faces, can be profoundly affected by contextual influences, at both encoding and retrieval (see Ceci & Bruck, 2006). It is thus critically important to identify factors that affect children's and adolescents' identification abilities.

In this study, we examined the effects of two potentially important factors—stress at encoding and interviewer-provided social support at retrieval—on children's and adolescents' lineup identification abilities. We were especially interested in whether stress at encoding interacted with interview context to affect children's accuracy, and whether this effect varied across age. Our results are relevant practically to evaluations of children's and adolescents' identification abilities and methods that may enhance those abilities in legal settings. Our results also have important theoretical implications for models of episodic memory development (see Bauer & Fivush, 2010), for debates about the effects of stress on memory (see Wolf, 2009), and for the notion of adolescence as a unique period of development in terms of stress reactivity and sensitivity to context (see Dahl & Gunnar, 2009; Flin, 1980).

In particular, extant theories of episodic and autobiographical memory development (e.g., Nelson & Fivush, 2004) have long emphasized that contextual factors at encoding influence which experiences become a part of one's autobiographical memories.

This research was supported by a grant from the National Science Foundation (BCS-0721377), and a grant-in-aid from the American Psychology-Law Society. We thank the families and children who participated in the project, as well as the undergraduate research assistants who assisted in data collection. Additional thanks to Jo Ann Prause for her help with the statistical analyses.

Correspondence concerning this article should be addressed to Jodi A. Quas, Department of Psychology and Social Behavior, 3340 Social Ecology Building II, University of California, Irvine, Irvine, CA 92697-7085. Electronic mail may be sent to jquas@uci.edu.

DOI: 10.1111/cdev.12177

The role of the context at retrieval (e.g., whether children are comfortable recounting a prior experience) has rarely been considered as an additional factor influencing autobiographical memory retention over time. As we discuss here, it may need to be studied directly. Also, an impressive literature has investigated, using rigorous experimental procedures, the effects of stress on adults' memory. No comparably large body of research has used similarly rigorous procedures to test the effects of stress on children's memory. It thus remains unclear as to whether findings in adults extend to younger individuals (Quas & Klemfuss, in press). Finally, the transition to adolescence is marked by acute increases in sensitivity to external stress, greater need for emotional regulatory abilities, and heightened attention to social interactions (Forbes & Dahl, 2010). All of these changes may influence adolescents' memory abilities, including for faces, specifically while under stress.

Developmental Changes in Identification Accuracy

Across studies of children's identification abilities, perhaps the most robust finding concerns age-related improvements in accuracy (Chance & Goldstein, 1984; Flin, 1980; Goldstein & Chance, 1964). For instance, in studies utilizing facial recognition paradigms, children view large numbers of faces and are then shown a second set of faces, some of which are repeated from the first set and some of which are new. When children are asked to identify faces they have seen previously, dramatic increases in hits and decreases in false alarms are evident for children between ages of 5 and 12 years (see Shapiro & Penrod, 1986, for a review).

Age-related improvements in performance are similarly evident in paradigms more closely resembling eyewitness identification scenarios. In these studies, children interact with an unfamiliar adult and later identify that adult via a lineup test. With lineups, of importance, age improvements emerge primarily in terms of lower false identifications when presented target-absent arrays (i.e., arrays that do not contain the target adult; Davies, Stevenson-Robb, & Flin, 1988; King & Yuille, 1987): Young children have considerable difficulty, and in fact, even adolescents do not perform as well as adults (Dekle, Beal, Elliot, & Honeycutt, 1996; Parker & Carranza, 1989; Parker & Ryan, 1993). It may be that children and adolescents feel increased pressure to pick someone from the lineup, even if he or she is not present (Pozzulo & Lindsay, 1997, 1998), perhaps because the context at retrieval (i.e., an array is presented) implies they should. Picking someone does help, though, when the target is actually present. Here, children as young as 5 perform virtually identically to older children and adults (e.g., Dekle et al., 1996; Pozzulo & Lindsay, 1998).

Effects of Stress on Memory and Identification Accuracy

Although studies of neurobiology and memory consistently find that stress at or near encoding enhances episodic memory in adults and in children (Quas, Rush, Yim, & Sumaroka, 2012; Schwabe, Joëls, Roozendaal, Wolf, & Oitzl, 2011), in eyewitness studies, high levels of stress are believed to impair identification accuracy (see Deffenbacher, Bornstein, Penrod, & McGorty, 2004; Morgan et al., 2004) in adults as well as children. In an influential meta-analysis, Deffenbacher et al. (2004) failed to uncover a significant moderating influence of witness age (dichotomized as child vs. adult) on the direction of observed effects across studies testing the effects of stress on lineup accuracy. Relatively few developmental studies, however, were actually included in the meta-analysis, and methodological variations across the studies preclude clear inferences about their findings. For instance, children's age varied considerably across studies, and age is strongly related to memory performance (e.g., Bruck & Ceci, 1999; Malloy & Quas, 2009). Collapsing analyses across age into a single group may have masked changes across childhood and adolescence in memory, identification abilities, or the ways in which stress affects mnemonic processes more generally (e.g., Dahl & Gunnar, 2009; Quas, Carrick, Alkon, & Boyce, 2006). Also, children were not always randomly assigned to high- and low-stress conditions; instead, differences in stress responses were correlated with memory (e.g., Goodman, Hirschman, Hepps, & Rudy, 1991). Factors other than stress (e.g., coping, attention) may have affected children's stress responses and memory. Finally, in some studies, children's age was negatively correlated with their stress responses (e.g., Peters, 1991), or the events in the high- and low-stress conditions varied (e.g., Goodman, Hirschman, et al., 1991; Peters, 1991). As such, it is difficult to draw conclusions from developmental studies, and it seems problematic to collapse children of all ages into a single group to be directly compared to adults.

In this study, we utilized a laboratory procedure that is comparably arousing across age, but amenable to experimental manipulation in the level of stress induced. We were thus able to test, in a rigorous manner, the effects of stress on children's identification abilities and determine whether the effects varied between children and adolescents.

Effects of Interviewer Behavior on Identification Accuracy

The context at retrieval has been well studied as a predictor of children's general memory and suggestibility. Typically, context is manipulated to be either positive and supportive or cold and intimidating (e.g., by varying whether interviewers behave supportively or not, whether children are interviewed in an intimidating courtroom or classroom setting, or whether interviewers build rapport or not; Bottoms, Quas, & Davis, 2007; Memorandum of Good Practice, see Bull, 1992; the National Institute of Child Health and Development Protocol, see Lamb, Orbach, Hershkowitz, Esplin, & Horowitz, 2007). Although supportive contexts do not consistently enhance the amount of information reported, they consistently decrease children's suggestibility (Almerigogna, Ost, Akehurst, & Fluck, 2008; Carter, Bottoms, & Levine, 1996; Davis & Bottoms, 2002; Goodman, Bottoms, Schwartz-Kenney, & Rudy, 1991). Supportive contexts may reduce children's anxiety, which in turn could allow them to focus more closely on an interviewer's questions before answering (Almerigogna, Ost, Bull, & Akehurst, 2007; Quas, Bauer, & Boyce, 2004). Children may also feel more comfortable disagreeing with suggestions provided by a warm rather than cold and potentially intimidating adult, a mechanism that has been termed "resistance efficacy" (Bottoms et al., 2007; Davis & Bottoms, 2002). To the extent that a supportive context helps children resist suggestions, it may be beneficial with lineup tasks, especially when the target is absent, given that these arrays are inherently suggestive in that children may feel pressure to choose someone, whether the perpetrator is present or not.

Although published research has not directly examined the effects of interviewer-provided social support on eyewitness identification accuracy, Lowenstein, Blank, and Sauer (2010) investigated 9- and 10-year-olds' ability to identify the perpetrator of a mock crime when the lineup administrator wore a uniform or casual clothes. Compared to children in the casual clothes condition, children in the uniform condition were more likely to make false identifications in target-absent lineups, were less likely to express uncertainty about their decision, and reported being more anxious. Insofar as a formal uniform creates an unsupportive context, heightened suggestibility, including identification errors, is not surprising.

Two important points are worth noting with regard to interview context effects. First, although findings generally show that high support reduces errors, the effects of support have not been compared across wide age ranges of children or examined in adolescents. Although suggestibility decreases with age, adolescents are certainly not immune to errors (e.g., Wright, London, & Waechter, 2010) and, as mentioned, false identification rates are higher in adolescents than in adults with target-absent lineups (Pozzulo & Lindsay, 1998). Thus, supportive contexts may moderate identification performance even in adolescents, although given children's greater deference to authority (Scherer, 1991), support effects may still be strongest in children. These potential developmental trends were tested directly in this study.

Second, and interestingly, supportive interviewing may moderate the effects of event stress on children's performance (see Bottoms et al., 2007). As mentioned, some work suggests that stress can enhance encoding, for example, by facilitating memory consolidation (see Roozendaal, 2002; Wolf, 2009), in contrast to the view that stress inhibits eyewitness identification. It is possible that some of the detrimental effects of stress on identification abilities may be due to stress at retrieval rather than encoding. Discussing a prior stressful event can elicit increased arousal or mild anxiety (Brenner, 2000; Levine, Burgess, & Laney, 2008), and stress at retrieval is associated with poorer memory, in children and adults (Kuhlman, Piel, & Wolf, 2005; Saywitz & Nathanson, 1993; Quas et al., 2004). Thus, stress experienced in the context of an interview may exacerbate effects of encoding stress on identification abilities. Insofar as supportive interviewing reduces arousal (Almerigogna et al., 2007), it may help children focus on and correctly respond to a lineup.

Although research has yet to test directly whether supportive interviewing mitigates, at least in part, the effects of stress on identification accuracy, Peters (1991) found some trends consistent with this possibility in a study of children's ability to identify the perpetrator of a mock crime. Children witnessed the crime either as naïve observers (high-stress encoding) or as a priori informed participants (low-stress encoding). The lineup was then conducted live or via a photo array, the latter of which may be considered more supportive, or at least less intimidating, because children were not in the physical presence of the perpetrator. Children in the high-stress encoding condition performed considerably better in the photo than live lineup condition (75% vs. 33% accuracy, respectively), as would be expected if the supportive context was influencing their reporting of a stressful experience. Children in the low-stress encoding condition were unaffected by lineup format (75% and 83% accuracy for the live and photo conditions, respectively). A direct test of the joint effects of encoding stress and interview support at retrieval, including across age, is clearly needed and was conducted here.

The Present Study

In this study, 7- to 8- and 12- to 14-year-olds completed either a high-stress (standard) or a low-stress version of the Trier Social Stress Test–Modified (TSST–M; Quas, Rush, Yim, & Nikolayev, 2013; Yim, Quas, Cahill, & Hayakawa, 2010; modified from the original TSST protocol, Kirschbaum, Pirke, & Hellhammer, 1993). In both the high- and low-stress versions of this structured laboratory protocol, individuals were required to complete a speech and arithmetic task in front of two unfamiliar adult observers, one male and one female. After a 2-week delay, participants returned for a memory interview during which their ability to accurately identify the observers was examined. Lineups were either target present or target absent, and interviews were conducted in a supportive or nonsupportive manner. The study conformed to a 2 (age: children vs. adolescents) × 2 (encoding stress: high vs. low) × 2 (retrieval context: supportive vs. nonsupportive) × 2 (lineup: target present vs. target absent) between-subjects design.

This design allowed us to evaluate children's and adolescents' lineup accuracy when recounting a stressful event and test one potential method—supportive interviewing—of improving that accuracy. The design also enabled us to take an important step toward advancing knowledge of changes in memory for faces and in the effects of stress on episodic memory across childhood and into adolescence. As mentioned, increasing evidence suggests that the transition to adolescence marks a unique period of development in terms of face recognition, stress reactivity, and sensitivity to context. As such, it is imperative to focus, in a complex way, on stress, context, and development as jointly contributing to memory, particularly facial identification accuracy.

Hypotheses included first that adolescents would correctly reject the target-absent lineups at a higher rate than children, consistent with prior research. Second, a main effect of interview condition was hypothesized, such that participants interviewed in a supportive manner would provide a greater number of correct identifications in the target-present and fewer errors (and a correspondingly higher number of correct rejections) in the target-absent lineup than participants interviewed in a nonsupportive manner. These latter effects were further expected to be subsumed by two significant interactions: Event Stress × Interview Support and Age × Interview Support. In particular, supportive interviewing was anticipated to be especially beneficial for participants who experienced the high-stress rather than low-stress TSST–M and for children relative to adolescents. Finally, because of the combination of children's greater deference to authority and greater suggestibility relative to adolescents, a three-way interaction among age, event stress, and interview support was expected, particularly for target-absent lineups, such that children in the high-stress condition would benefit from a supportive interview to a greater extent than adolescents in the high-stress condition.

Method

Participants

Families were recruited via telephone by a market research firm (Fieldwork, Los Angeles, CA) or by an existing database of families interested in research. Families with children in the eligible age ranges were contacted and, if interested, scheduled for the study. All participants were fluent in English and free from serious medical or mental health problems. The final sample in the current report contained 78 children (7–8 years, $M = 7.5$, $SD = .50$, 48.7% males) and 81 adolescents (12–14 years, $M = 12.98$, $SD = .88$, 48.1% males). Ethnicity varied and included 53% European American, 8% Hispanic/Latino, 4% African American, 2% Asian, 30% multiethnic, and 3% other, and parents were fairly highly educated: 47% of mothers and 56% of fathers had a bachelor's degree or higher. Finally, 79% of parents reported a household annual income of $60,000 or more. Ten additional participants (six children and four adolescents) were excluded due to lineup administration errors ($n = 7$), failure to complete the entire TSST–M ($n = 2$), or failure to return for the second session ($n = 1$). They did not differ from included participants on age, gender, or experimental conditions (ts = −.93–.43, ps = .37–.74).

Materials and Procedures

Session 1

Upon arrival, a researcher explained the study in detail to parents and obtained their consent. The researcher then described the session to participants and obtained their assent. The descriptions did not include mention of the memory or identification task. Following consent, parents completed a demographic and health screening questionnaire, and participants were escorted to a quiet room for a 20-min rest period while they completed unrelated questionnaires (e.g., their mood).

Participants were randomly assigned to either the standard TSST–M, which is moderately to highly stressful, or the low-stress TSST–M. For heuristic purposes, we refer to the two conditions as "high-stress" and "low-stress" (see Quas et al., 2013, for details). Because the observers who administered the TSST–M behaved differently in the high- and low-stress TSST–M conditions, and because the researcher giving the TSST–M overview instructions needed to vary her instructions between conditions, blinding of research assistants (RAs) during the TSST–M phase was neither possible nor desirable.

The TSST-M occurred in a separate room. When participants entered, a male and a female observer in white lab coats were waiting. The researcher then explained the TSST–M procedure. The general instructions for the TSST–M were essentially the same between the two stress conditions: The researcher told participants that they would be asked to give a 5-min speech and complete a mathematics task. She further explained that during the speech, participants should pretend that they are in a new class, introduce themselves, and tell the class about their personality; and that the mathematics task would start after the speech. Beyond the comparability of instructions, however, were several important differences between the high- and low-stress conditions (see Quas et al., 2013).

In the high-stress TSST–M, the researcher created an evaluative context (e.g., participants were told that they would be videotaped, and that the videotapes would be analyzed later). Participants had a 3-min preparation period during which the observers remained silent and maintained a neutral demeanor. They kept the same demeanor throughout the tasks as well. The male observer administered the 5-min speech task instructions and, if participants stopped talking before the time allotted, he told them that they still had more time and then asked scripted questions in 30-s intervals (e.g., "Tell us the three best things about you," "Do you like being part of a team?"). The female observer administered the 5-min mathematics task, which required participants serially subtract a number from a bigger number out loud (the subtraction numbers and starting points varied based on participants' grade level). She corrected errors and had participants start over from the beginning each time. These procedures are similar to those used in numerous prior TSST studies with adults and children (e.g., Gordis, Granger, Susman, & Trickett, 2006; Kirschbaum et al., 1993; Yim et al., 2010).

In the low-stress condition, participants still had a 3-min preparation period and completed the same speech and mathematics task. Modifications were introduced to make the context of the TSST–M less evaluative or test like and more positive interpersonally. Participants were informed that they would be videotaped, but further told that this was only as a backup. The observers were introduced and said hello at the start of the preparation period. They maintained a positive demeanor throughout the tasks. The researcher also said that the male observer was new, was still training, and may make mistakes. If participants stopped talking before 5 min had ended on the speech task, the male observer encouraged them to tell more and then asked the scripted questions in 30-s intervals, but he fumbled through writing down information to avoid periods of awkward silence. Likewise, during the mathematics task, the female observer still corrected participants and had them start again, but she said that this was so she could follow along.

At the end of the TSST–M, participants were thanked and told that they did a great job. They were taken back to another room where they completed a questionnaire with nine items concerning their perceptions of their performance and feelings during the TSST–M. All items were rated on a 7-point Likert-type scale. Sample items included: "This session was …" (1 = *not at all hard*, 7 = *extremely hard*); "When I did the speech, I thought I did …" (1 = *very poorly*, 7 = *very well*); "When I did the speech, I felt …" (1 = *not at all stressed*, 7 = *very stressed*). This questionnaire was included as a manipulation check to assess whether participants indeed responded differently to the high-stress and low-stress versions of the TSST–M. Participants then completed questionnaires unrelated to this study. Afterward, participants were told that they would be coming back for a visit in 2 weeks, but that the next visit would take place in a separate building and would consist of different activities. They were not informed of our interest in memory.

Session 2

Session 2 took place after a 2-week delay (M = 14 days, SD = 1.68 days). The TSST–M observers were not present, and interviewers were naïve to participants' individual TSST–M experiences. After parents approved the materials, participants were taken to the interview room where an unfamiliar female interviewer was waiting. She administered the interview, which included the lineup identification task, in either a supportive or nonsupportive manner, using procedures employed in prior studies of interviewer support and children's eyewitness abilities (e.g., Carter et al., 1996; Davis & Bottoms, 2002; Quas & Lench, 2007). Children in each age and TSST–M condition were randomly assigned to interview condition.

In the supportive condition, the interviewer dressed casually and introduced herself. During the interview, she maintained eye contact, smiled frequently, and provided positive verbal feedback at designated times. In the nonsupportive condition, the interviewer dressed professionally and did not introduce herself. During the interview, she maintained a neutral demeanor, and did not smile or provide positive feedback. She also informed participants that there were observers behind a mirror evaluating the participants. At specified times, she asked participants to speak up so that the observers could hear them.

The interview concerned what happened during the entire prior session (e.g., "Tell me everything that happened when you went to the other building and completed the first part of the study"; Quas et al., 2013). At the end of the interview, the interviewer administered two 6-person photographic simultaneous lineups via a standard lineup procedure task while maintaining the same supportive or nonsupportive demeanor used in the rest of the interview. She began by reading scripted instructions stating participants would view photographs that may or may not contain a picture of the observers from the last session. The male lineups were presented first. She instructed participants to state the number under the picture of the male observer if he was in the array, say "not there" if his picture was not there, or say "I don't know" if they were not sure. She then presented the lineup array, which showed all photos simultaneously so participants could compare them directly. Half of the participants viewed a target-present lineup, and half viewed a target-absent lineup. Lineup type was counterbalanced across age (child vs. adolescent), sex, TSST–M stress condition (high vs. low), and interview condition (supportive vs. nonsupportive). The array was presented until the child gave a response (no child failed to give any response). The interviewer recorded the response without giving any feedback and then repeated the process for the female observer lineup.

The male lineup was always shown first to match the temporal order of participants' experiences (participants interacted with him first and did so more extensively than with the female observer). In addition, males are substantially more likely than females to be suspects of crime. Thus, we wanted to ensure that the participants had not had any practice with a lineup before viewing the male lineup and we did not want order effects on their performance with the male lineup. After administering the female lineup, the interviewer thanked participants and left.

As the interviewers and observers were acquainted, interviewers were instructed not to look at the lineups and were not told whether the lineup was target present or target absent. This minimized the chances of the interviewer communicating information about which photo to choose. Interviewers then retrieved the lineup from a closed folder and held it in front of themselves at arm's length, facing the participant. Interview videotapes confirmed that the interviewers followed the instructions.

Following the interview and identification procedure, another experimenter administered a final questionnaire, which included three items about how stressful the interview was and, for a subset of participants (n = 68), items asking about participants' perceptions of the interviewer, all rated on a 7-point Likert-type scale. Sample items included "During the interview, the interviewer was …" (1 = *not at all nice*, 7 = *very nice*), "During the interview, the interviewer was …" (1 = *not at all serious*, 7 = *very serious*), and "During the interview, I felt …." (1 = *not at all stressed*, 7 = *very stressed*). These questions were included as manipulation checks regarding the interviewer behavior. After the session, participants were fully debriefed, which included informing participants that the first session tasks were meant to be very challenging and that they did an excellent job, and for participants in the nonsupportive condition that the interviewer was instructed to behave in a neutral manner. Families were thanked and paid a small honorarium.

Lineup Construction

The procedures for creating lineups followed those used most frequently by law enforcement (Wise, Safter, & Maro, 2011; Wogalter, Malpass, & McQuiston, 2004). Target-present and target-absent lineups were constructed for each male (n = 7) and female (n = 11) RA who served as TSST–M observers, producing 36 lineups. Basic descriptive information for the RAs (i.e., ethnicity, height, weight, hair, and eye color) was entered into the Florida Department of Corrections database of mug shot photographs (http://www.dc.state.fl.us/) to generate a pool of 100–200 photographs for each RA. A panel of independent researchers, unaware of the study hypotheses, constructed the target-present lineups by comparing each RA (target) photograph to the pool of potential matches to select six photos that matched in terms of a general verbal description and visual appearance. Then, for each RA target, one of the six selected photographs was randomly designated to be the innocent suspect in the target-absent lineup and the other five were used as foils in the target-present lineup.

The target-absent lineups were constructed by a different panel of researchers that had no exposure to the targets and no involvement in the construction of target-present lineups (all were also naïve to the study design and hypotheses). The procedure for constructing target-absent lineups was identical to that for the target-present lineups. For each target-absent lineup, six foils were selected that resembled the innocent suspect (from the same pool of photographs used for the target-present lineup). One of the six foils was randomly set aside and not used in any lineup, and the other five were designated as foils in the target-absent lineup.

Once the suspect and foils were finalized, photos were digitally altered for homogeneity: A uniform shade was used, brightness was made consistent, backgrounds and clothing were made identical, and individual markings and distinguishing features (e.g., tattoos) were removed.

Coding

We calculated four accuracy variables for each participant. First, we created a binary accuracy variable that simply reflected whether the participants' lineup decision was correct or incorrect. Second, we created a categorical variable that reflected all possible lineup decision outcomes. Thus, for target-present lineups, responses could include correct identification, foil (false) identification, incorrect rejection of the lineup, or "don't know." For target-absent lineups, responses could include correct rejection of the lineup, false identification of innocent suspect, foil identification, or "don't know." These two accuracy variables were calculated for both the male and female lineups for each participant. We also created a binary choosing variable reflecting whether or not participants chose someone out of each lineup, regardless of accuracy.

Results

Our between-subjects study design was as follows: 2 (age: 7–8 vs. 12–14 years) × 2 (TSST–M stress: high vs. low) × 2 (interview context: supportive vs. nonsupportive) × 2 (lineup type: target present vs. target absent). Rates of correct and incorrect identification decisions, as well as choosing rates, across age, TSST–M condition, and interview condition for the male observer are presented in Table 1 (*ns* in parentheses). It is important to note that in real-world target-absent lineups, false identification of an innocent suspect is more forensically consequential than the selection of a foil from the same lineup (Wells & Turtle, 1986). However, because of small cell sizes in this study, we could not examine innocent suspect ID rates separately from foil ID rates, and thus both of these errors are considered "false" identifications in the analyses.

Analyses performed for the female lineups revealed no significant effects. As mentioned, however, the female lineup was always shown second, presenting a confound that precludes us from interpreting the reason for the nonsignificant findings. Thus, the female lineup data are not considered further.

Preliminary Analyses and Manipulation Checks

First, the effects of gender, ethnicity, observer, and delay were examined. No significant differences in identification accuracy (correct vs. incorrect) emerged for participants' gender or ethnicity, χ^2s(1, 6) ranged from .065 to 10.69, ps > .06. Which male RA administered the TSST–M (n = 7) was also unrelated to participants' identification accuracy, $\chi^2(6) = 4.193$, p = .65, as was the delay between sessions, r = .03, p = .68. None of these variables is considered further.

Second, we assessed the effectiveness of our TSST–M and interviewer support manipulations.

Table 1
Percentage (N) of Correct and Incorrect Identification Decisions by Age and Condition

	Children				Adolescents			
	High stress		Low stress		High stress		Low stress	
Decision	Nonsupportive	Supportive	Nonsupportive	Supportive	Nonsupportive	Supportive	Nonsupportive	Supportive
Target present (N = 83)								
Correct ID	60 (6)	55 (6)	55 (5)	50 (5)	40 (4)	64 (7)	91 (10)	54 (6)
Foil ID	20 (2)	36 (4)	11 (1)	0 (0)	30 (3)	0 (0)	0 (0)	18 (2)
False rejection	20 (2)	9 (1)	11 (1)	40 (4)	20 (2)	27 (3)	9 (1)	18 (2)
Don't know	0 (0)	0 (0)	22 (2)	10 (1)	10 (1)	9 (1)	0 (0)	9 (1)
ID rate[a]	80 (8)	91 (10)	67 (7)	50 (5)	70 (7)	64 (7)	91 (10)	73 (8)
Target absent (N = 76)								
Correct rejection	13 (1)	70 (7)	36 (4)	11 (1)	60 (6)	70 (7)	56 (5)	33 (3)
Foil ID	50 (4)	0 (0)	36 (4)	33 (3)	10 (1)	10 (1)	33 (3)	67 (6)
False ID	13 (1)	10 (1)	0 (0)	11 (1)	0 (0)	20 (2)	11 (1)	0 (0)
Don't know	25 (2)	20 (2)	27 (3)	44 (4)	30 (3)	0 (0)	0 (0)	0 (0)
ID rate[a]	63 (5)	10 (1)	36 (4)	44 (4)	10 (1)	30 (3)	44 (4)	67 (6)

[a]Percent who chose someone in the lineup.

Participants' responses to four items on the post-TSST questionnaire that asked how challenging the TSST–M was overall, how stressful participants found the speech and mathematics components to be, and to what extent they wanted to stop were averaged (rs between the items ranged from .32 to .56, ps < .001, α = .73) such that higher scores reflect more perceived stress. Participants in the high-stress (standard TSST–M) condition reported higher levels of stress, M = 4.02, SD = 1.36, than participants in the low-stress condition, M = 3.39, SD = 1.25, $t(157) = -3.09$, $p < .01$. Physiological arousal data collected as a part of the larger project confirmed the self-report data: Participants in the high-stress TSST–M condition evidenced larger hypothalamic pituitary adrenal axis and sympathetic nervous system responses than participants in the low-stress TSST–M condition (see Quas et al., 2013, for details). Of importance, analyses of the stress data indicate that the low-stress TSST–M was not *non*stressful, as participants' self-report ratings were near the midpoint of the scale. However, we refer to the two conditions as high and low for readability and ease of interpretation.

Participants' responses to two post-interview items about the interviewer's demeanor were also averaged (items reverse coded so that higher scores indicate more positive as appropriate). Participants rated the supportive interviewer more positively, M = 4.53, SD = .1.39, than the nonsupportive interviewer, M = 3.60, SD = 1.52, $t(66) = 2.75$, $p < .01$. In addition, participants' responses to questions about their reaction to the interview (e.g., how stressfulness it was) were averaged (rs between the items ranged from .32 to .57, ps < .01, α = .73). Participants in the nonsupportive condition rated the interview more stressful, M = 3.01, SD = 1.43, than participants in the supportive condition, M = 2.54, SD = 1.23, $t(156) = -2.20$, $p < .05$. (Participants' physiological responses to the interviews did not differ significantly; see Quas et al., 2013.)

Identification Accuracy

Our primary interest concerned the accuracy of children's and adolescents' lineup identifications. We tested the study's hypotheses via logistic regressions predicting identification accuracy. We first constructed a main effects model to address the independent contributions of age, TSST–M, and interviewer support. We then conducted separate models testing the two-way interactions, followed by a model including the higher order three-way interaction. We present coefficients only for the models that include the highest order of significant interaction terms. For significant interactions, follow-up analyses with 95% confidence intervals (CI) for odds ratios (OR) identified significant contrasts for variables in the interaction (Hosmer & Lemeshow, 2000).

Target-Present Lineups

First, we examined identification decisions of participants who viewed target-present lineups (n = 83), with accuracy dichotomized into "correct" (identification of the observer, or hit) and "not correct" (filler identification, false rejection, and "don't know" responses). None of the models was significant (χ^2s = 1.49–7.75, ps = .35–.68). Thus, age, event stress, and interview context did not significantly predict participants' lineup identification accuracy when the lineups contained the target photograph.

Target-Absent Lineups

Next, we repeated the analyses predicting decisions of participants who viewed target-absent lineups (n = 76). The final model was significant, $\chi^2(4) = 13.51$, $p < .01$, and included the main effects of age, TSST–M condition, and interview condition, as well as the interaction of TSST–M condition and interview condition. The main effect of age was significant, Wald(1) = 3.89, $p < .05$, OR = 1.67. As expected, adolescents had a higher percentage of correct rejections (55.3%) than children (34.2%). The main effect of TSST–M condition was also significant, Wald(1) = 8.22, $p < .01$, OR = 9.25, but was subsumed by a significant TSST–M Condition × Interview Condition interaction, Wald(1) = 6.21, $p < .05$ (see Figure 1). Consistent with our hypothesis, the highest proportion of correct rejections (70%) was evident among participants who experienced the high-stress TSST–M and were interviewed in a supportive manner. Contrast analyses revealed that the odds of a correct rejection varied significantly based on interview condition when TSST–M condition was high stress (95% CI for the OR: −2.83, −.03): Participants in the high-stress TSST–M condition were significantly more likely to correctly reject the target-absent lineup when interviewed in a supportive (70%) rather than nonsupportive (39%) manner. In addition, the odds of a correct rejection varied significantly based on TSST–M condition when participants were interviewed supportively (95% CI for the OR: 2.26, 37.05): Participants interviewed in a supportive manner were better able to reject the lineup if they had experienced the high- rather than low-stress TSST–M. Examination of the correct rejection rates (see Table 1) revealed that the interaction was largely driven by the children (although the three-way interaction was not significant, most likely because of a lack of power). Correct rejection rates for children in the high-stress TSST–M condition increased from 13% to 70% across the nonsupportive to supportive interview conditions. For adolescents, the correct rejection rate increased less dramatically (60% to 70%).

Of note, the dependent measure in the target-absent analyses collapsed false identifications, true errors, and "don't know" responses into a single, "not correct" group. Given that "don't know" responses, particularly in terms of legal relevance, are fundamentally different from false identifications, it was important to test systematically the effects of stress and interviewer support on the two types of responses. Thus, a multinomial logistic regression was carried out using a dependent variable composed of the three possible target-absent lineup decisions: correct rejection, false identification, and "don't know," with correct rejections serving as the reference group. The final model was again significant, $\chi^2(8) = 18.21$ $p < .05$.

For false identifications, the TSST–M Condition × Interview Condition interaction was significant, Wald(1) = 4.16, $p < .05$ (see Figure 2). Consistent with the results for correct rejections, the fewest number of false identifications (20%) occurred among participants who experienced the high-stress TSST–M and were then interviewed in a supportive manner. Contrast analyses revealed that among participants in the supportive interview condition, those who experienced the high-stress TSST–M were less likely to make a false identification (20%) than those who experienced the low-stress TSST–M (56%; 95% CI for the OR: .02, .55). Inspection of the false ID rates (see Table 1) revealed that the interaction was again driven largely by the younger children (although, and also likely as a result of power, the three-way interaction was nonsignificant). Among children in the high-stress TSST–M condition, false identifications fell from

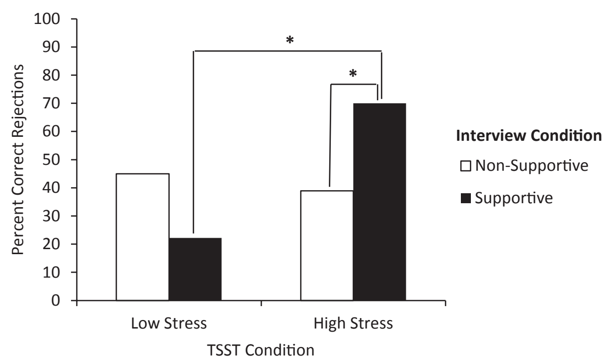

Figure 1. Percentage of correct rejections in target-absent lineups by stress and interviewer support condition. TSST = Trier Social Stress Test; * = significant differences per contrast effects.

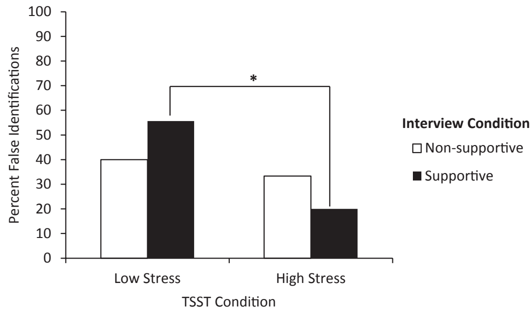

Figure 2. Percentages of false identifications in target-absent lineups by stress condition and interviewer support condition. TSST = Trier Social Stress Test; * = significant differences per contrast effects.

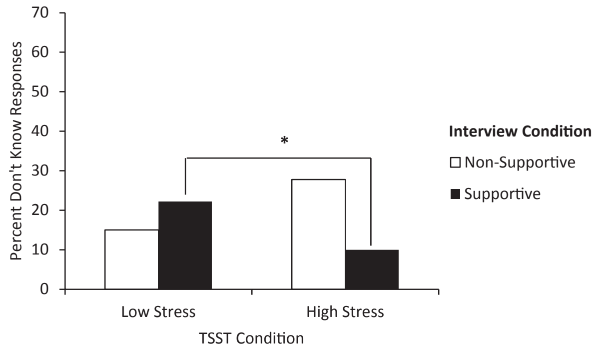

Figure 3. Percentages of "don't know" responses in target-absent lineups by stress condition and interviewer support condition. TSST = Trier Social Stress Test; * = significant differences per contrast effects.

63% to 10% across the nonsupportive to supportive conditions. In contrast, among adolescents in the high-stress TSST–M condition, false identifications were actually slightly higher in the supportive (30%) than in the nonsupportive (10%) condition.

For "don't know" responses, the effect of age was significant, Wald(1) = 6.51, $p < .05$, OR = .37. Children (29%) were more likely to say "I don't know" than adolescents (8%). Again, the TSST–M Condition × Interview Condition interaction was significant, Wald(1) = 4.87, $p < .05$ (see Figure 3). Participants were least likely to say "I don't know" when they experienced the high-stress TSST–M and subsequently the supportive interview (10%). Contrast analyses revealed that among participants in the supportive interview condition, those who experienced the high- rather than low-stress TSST–M were less likely to answer "I don't know" (10% vs. 22%, respectively), 95% CI for the OR: −4.27, .002, again consistent with our hypotheses.

In a final set of analyses, we assessed participants' likelihood of choosing anyone from the lineup rather than rejecting the lineup or saying "don't know" as a function of interviewer support (Table 1). Lowenstein et al. (2010) found that children interviewed by a uniformed (and hence perhaps more intimidating) rather than plain clothed officer were more likely to choose, which could explain the increase in false identifications in target-absent lineups. Chi-square analyses comparing choosing rates across interviewer condition, however, failed to reveal any significant effects for children, $\chi^2(1) = .87$, $p = .35$, or adolescents, $\chi^2(1) = .10$, $p = .75$.

Discussion

The purpose of this study was to examine, using controlled, rigorous laboratory procedures, the effects of stress and interview context on children's and adolescents' lineup identification accuracy. Two key sets of results emerged. One, which has been repeatedly demonstrated, was that of age differences in performance. The other, which is more novel and practically and theoretically informative, was that of the combined effects of encoding stress and retrieval context on children's and adolescents' eyewitness identification abilities.

First, on the basis of previous findings (Pozzulo & Lindsay, 1998), we predicted—and found—age-related improvements in identification accuracy, but only for target-absent lineups. When the confederate was included in the lineup, children's and adolescents' performance was virtually identical (55% and 62%, respectively). When the confederate was absent, children had much more difficulty (correct rejections were 34% for children and 55% for adolescents). Thus, the typically observed age-related increases in performance do not appear altered when the to-be-remembered event is stressful. That being said, nearly 50% of the adolescents made incorrect decisions when the target was absent, suggesting, as other developmental studies have, that the ability to correctly reject target-absent lineups continues to develop well into adolescence (Pozzulo & Lindsay, 1998).

Second, the hypothesized interaction between encoding stress and retrieval stress emerged: Participants who experienced the high-stress TSST–M were especially likely to benefit, in terms of rejecting target-absent lineups, from a socially supportive interviewer. Recounting a stressful event can induce some anxiety or arousal (Brenner, 2000; Levine et al., 2008), and arousal at retrieval is consistently associated with lower memory accuracy in children (Almerigogna et al., 2007; Nathanson & Saywitz, 2003; Quas & Lench, 2007; Saywitz & Nathanson, 1993), a pattern often attributed to their need to devote resources during an interview toward regulating their arousal rather than toward conducting an adequate memory search (Quas et al., 2004). The supportive interviewer reduced children's anxiety, per the self-report data. However, support was only beneficial when participants were evaluating the target-absent lineups. It did not influence participants' general memory for the male observer as reflected in their performance on the target-present lineup. Thus, interviewer support seems to have influenced participants' comfort rejecting the target-absent lineup, and not their ability to conduct an adequate memory search. When the original event was stressful, participants (particularly children) may have needed extra reassurances to feel comfortable stating that the confederate was not present. Also, self-efficacy has been suggested as a mechanism by which supportive interviewing helps children resist misleading information (Davis & Bottoms, 2002). In our study, participants who experienced the high-stress TSST–M may have come into the interview with lower feelings of self-efficacy, possibly as a result of the social evaluative components embedded in the TSST–M. The supportive interviewer may have increased these participants' confidence and willingness to reject the target-absent lineup. Which possibility, or combination thereof, actually drove the effect is not clear, but is certainly of interest for future research.

In addition, and to our surprise, when the TSST–M was low stress, it appeared that supportive interviewing may have *negatively* affected performance on target-absent lineups. We are uncertain of the reason for this pattern, although perhaps participants who experienced the low-stress TSST–M and supportive interview did not take the identification task as seriously as the other participants (and thus correct rejection rates were higher for participants interviewed in a more serious manner). Given that other investigations of supportive interviewing have not observed any detrimental effects, it will be important to replicate these findings in future lineup research. One can easily point to examples, for instance, of situations where children are victimized or are witness to a crime in an otherwise potentially positive or at least nonstressful context (as when an abuser is a trusted family member or friend). Although these situations would not likely be the types of cases in which a lineup identification would be conducted, it will be critical to evaluate whether supportive interviewing inhibits performance on other memory outcomes when recounting positive experiences, which could be relevant to the latter types of cases (see Klemfuss, Milojevich, Yim, Rush, & Quas, 2013, for evidence of support affecting the content of what is included in children's narratives).

One other point is worth noting regarding these data. Although age (children vs. adolescents) did not further moderate the observed interactive between event stress and interview context predicting false identifications, the overall pattern hints at developmental differences (i.e., with the findings appearing much stronger in the children, as we had hypothesized). Given that adolescents' performance may have followed a somewhat different pattern, and given other research indicating that early adolescence (i.e., puberty onset) is a time of increased stress reactivity and sensitivity to context (Dahl & Gunnar, 2009; Forbes & Dahl, 2010), exploration of the effects of supportive interviewing across the entire period of adolescence is warranted.

Finally, and notably, we did not find overall detrimental effects of event stress on participants' lineup identification accuracy, as reported in previous research (e.g., Deffenbacher et al., 2004; Morgan et al., 2004). Some nonsignificant hints at detrimental effects of stress emerged for target-present lineups, but the pattern actually appeared reversed in target-absent lineups. Overall, our results highlight the need for greater attention in both theories of neurobiological stress responses and memory (e.g., Wolf, 2009), and theories of eyewitness processes (Deffenbacher et al., 2004), to the complex ways in which stress at encoding and retrieval may jointly influence different types of memory and reporting processes, including those that require recall and identification of faces.

Limitations to the study also deserve mention. For one, the TSST–M reliably induces stress responses across age (see Quas et al., 2013; Yim et al., 2010) and confers clear advantages in terms of reliable stress manipulation compared to other stressors (e.g., fire alarms) used with children in past research. However, stress responses to the TSST–M likely do not approximate those experienced by witnesses and victims of extremely stressful, violent crimes. Thus, we cannot rule out that a substantial drop in accuracy would not occur at substantially higher levels of stress. Related, the TSST–M involves an element of inward focus, in that the stressful nature of the event is tied to its evaluative components, which is potentially in contrast to some real-world stressors, particularly crimes, for which the focus of the event is on a perpetrator's actions. Level of personal involvement and focus, in conjunction with age and stress, should be examined in more detail. Next, because of our relatively small sample size and low rates of true false identifications (i.e., identification of the designated innocent suspect from the target-absent lineup), we were not able to analyze these responses separately, electing instead to combine foil identifications and innocent suspect identifications into one "false ID" category. As mentioned, in actual investigation lineups, foils are "known innocents." The forensic consequences of identifying a foil are far less significant than those of identifying an innocent suspect (Wells & Turtle, 1986). Future research utilizing larger samples is needed to examine effects of stress and interviewer-provided social support on innocent suspect identifications specifically. Finally, because of timing limitations, we were not able to pilot test the target-present and -absent lineups for fairness and comparability. However, the procedure used to create the lineups, specifically the selection of foils based (in part) on their match to the suspect, follows the procedure that is most frequently used by law enforcement (Wise et al., 2011; Wogalter et al., 2004). Suspect-matched foil selection creates lineups that are often biased (Clark & Tunnicliff, 2001), an outcome that has emerged in actual criminal cases (Brigham, Meissner, & Wasserman, 1999). Indeed, our target-absent lineups showed precisely this pattern. Summing across conditions, 6 witnesses identified the innocent suspect and 22 identified a foil. Thus, the likelihood of identifying the innocent suspect (given that an identification was made), 6/28 = .21, was slightly higher than chance, consistent with real-world data.

In conclusion, this study provides much-needed insight into the relations among stress at encoding, retrieval context, and eyewitness identification abilities in children and adolescents. Prior studies have painted a very positive picture of the benefits of supportive interviewing, at least for reducing suggestibility, with no concurrent negative effects. Our study confirmed results of former studies, again demonstrating improved performance, in this case on target-absent lineups when even the original was highly stressful. However, hints at detrimental effects on target-absent lineup performance emerged when the original event was less stressful. Although the small cell sizes preclude our ability to generate strong recommendations based on the latter trend, these intriguing interactive results highlight the need for greater attention in future research to potentially complex associations among encoding stress and interview context in relation to children's and adolescents' memory and suggestibility. It is only with more extensive studies that clear recommendations of direct relevance to the legal system about methods of facilitating children's and adolescents' identification abilities can be generated, and more informed, broader theories of stress and episodic memory including for faces can be formulated.

References

Almerigogna, J., Ost, J., Akehurst, L., & Fluck, M. (2008). How interviewers' nonverbal behaviors can affect children's perceptions and suggestibility. *Journal of Experimental Child Psychology*, *100*, 17–39. doi:10.1016/j.jecp.2008.01.006

Almerigogna, J., Ost, J., Bull, R., & Akehurst, L. (2007). A state of high anxiety: How nonsupportive interviewers can increase the suggestibility of child witnesses. *Applied Cognitive Psychology*, *21*, 963–974. doi:10.1002/acp.1311

Bauer, P. J., & Fivush, R. (2010). Context and consequences of autobiographical memory development. *Cognitive Development*, *25*, 303–308. doi:10.1016/j.cogdev.2010.08.001

Bottoms, B. L., Quas, J. A., & Davis, S. L. (2007). The influence of interviewer provided social support on children's suggestibility, memory, and disclosures. In M.-E. Pipe, M. E. Lamb, Y. Orbach, & A.-C. Cederborg (Eds.), *Child sexual abuse: Disclosure, delay, & denial* (pp. 135–158). Mahwah, NJ: Erlbaum.

Brenner, E. (2000). Mood induction in children: Methodological issues and clinical implications. *Review of General Psychology*, *4*, 264–283. doi:10.1037/1089-2680.4.3.264

Brigham, J. C., Meissner, C. A., & Wasserman, A. W. (1999). Applied issues in the construction and expert assessment of photo lineups. *Applied Cognitive Psychology*, *13*, S73–S92 doi:10.1002/(SICI)1099-0720(199911)13:1+<S73::AID-ACP631>3.0.CO;2-4.

Bruck, M., & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual Review of Psychology*, *50*, 419–439. doi:10.1146/annurev.psych.50.1.419

Bull, R. (1992). Obtaining evidence expertly: The reliability of interviews with child witnesses. *Expert Evidence: The International Digest of Human Behaviour, Science and Law*, *1*, 3–36.

Bureau of Justice Statistics. (2008). *Trends in violent victimization by age: 1973–2008*. Retrieved from http://bjs.ojp.usdoj.gov/content/glance/vage.cfm

Carter, C. A., Bottoms, B. L., & Levine, M. (1996). Linguistic and socioemotional influences on the accuracy of children's reports. *Law and Human Behavior*, *20*, 335–358. doi:10.1007/BF01499027

Ceci, S. J., & Bruck, M. (2006). Childrens' suggestibility: Characteristics and mechanisms. *Advances in Child Development and Behavior*, *34*, 247–281. doi:10.1016/S0065-2407(06)80009-1

Chance, J. E., & Goldstein, A. G. (1984). Face recognition memory: Implications for children's eyewitness testimony. *Journal of Social Issues*, *40*, 69–85. doi:10.1111/j.1540-4560.1984.tb01094.x

Clark, S. E. (2012). Costs and benefits of eyewitness identification reform. *Psychological Science and Public Policy*, *7*, 238–259. doi:10.1177/1745691612439584

Clark, S. E., & Tunnicliff, J. L. (2001). Selecting lineup foils in eyewitness identification experiments: Experimental control and real-world simulation. *Law and Human Behavior*, *25*, 199–216. doi:10.1023/A:1010753809988

Cutler, B. L., & Penrod, S. (1995). *Mistaken identification: The eyewitness, psychology, and the law*. Cambridge, UK: Cambridge University Press.

Dahl, R. E., & Gunnar, M. R. (2009). Heightened stress responsiveness and emotional reactivity during pubertal maturation: Implications for psychopathology. *Development and Psychopathology*, *21*, 1–6. doi:10.1017/S0954579409000017

Davies, G. M., Stevenson-Robb, Y., & Flin, R. (1988). Tales out of school: Children's memory for an unexpected event. In: M. Gruneberg, P. Morris, & R. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 1, pp. 122–127). Chichester, UK: Wiley.

Davis, S. L., & Bottoms, B. L. (2002). Effects of social support on children's eyewitness reports: A test of the underlying mechanism. *Law and Human Behavior*, *26*, 185–215. doi:10.1023/A:1014692009941

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. doi:10.1007/s10979-004-0565-x

Dekle, D. J., Beal, C. R., Elliot, R., & Honeycutt, D. (1996). Children as witnesses: A comparison of lineup versus showup identification methods. *Applied Cognitive Psychology*, *10*, 1–12. doi:10.2139/ssrn.1443544

Flin, R. H. (1980). Age effects in children's memory for unfamiliar faces. *Developmental Psychology*, *16*, 373–374. doi:10.1037/0012-1649.16.4.373

Forbes, E. E., & Dahl, R. E. (2010). Pubertal development and behavior: Hormonal activation of social and motivational tendencies. *Brain and Cognition*, *72*, 66–72. doi:10.1016/j.bandc.2009.10.007

Goldstein, A. G., & Chance, J. E. (1964). Recognition of children's faces. *Child Development*, *35*, 129–136.

Goodman, G. S., Bottoms, B. L., Schwartz-Kenney, B. M., & Rudy, L. (1991). Children's testimony about a stressful event: Improving children's reports. *Journal of Narrative and Life History*, *1*, 69–99.

Goodman, G. S., Hirschman, J. E., Hepps, D., & Rudy, L. (1991). Children's memory for stressful events. *Merrill-Palmer Quarterly*, *37*, 109–157.

Gordis, E. B., Granger, D. A., Susman, E. J., & Trickett, P. K. (2006). Asymmetry between salivary cortisol and α-amylase reactivity to stress: Relation to aggressive behavior in adolescents. *Psychoneuroendocrinology*, *31*, 976–987. doi:10.1016/j.psyneuen.2006.05.010

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). Hoboken, NJ: Wiley.

King, M. A., & Yuille, J. C. (1987). Suggestibility and the child witness. In: S. G. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children's eyewitness memory* (pp. 24–35). New York, NY: Springer-Verlag.

Kirschbaum, C., Pirke, K., & Hellhammer, D. H. (1993). The "Trier Social Stress Test": A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*, 76–81. doi:10.1159/000119004

Klemfuss, J. Z., Milojevich, H. M., Yim, I. S., Rush, E., & Quas, J. A. (2013). Stress at encoding, context at retrieval, and children's narrative content. *Journal of Experimental Child Psychology*, Advance online publication. doi:10.1016/j.jecp.2013.07.009

Kuhlman, S., Piel, M., & Wolf, O. (2005). Impaired memory retrieval after psychosocial stress in healthy young men. *Journal of Neuroscience*, *25*, 2977–2982. doi:10.1523/JNEUROSCI.5139-04.2005

Lamb, M. E., Orbach, Y., Hershkowitz, I., Esplin, P. W., & Horowitz, D. (2007). A structured forensic interviewing protocol improves the quality and informativeness of investigative interviews with children: A review of research using the NICHD protocol. *Child Abuse and Neglect*, *31*, 1201–1231. doi:10.1016/j.chiabu.2007.03.021

Levine, L. J., Burgess, S. L., & Laney, C. (2008). Effects of discrete emotions on young children's suggestibility. *Developmental Psychology*, *44*, 681–694. doi:10.1037/0012-1649.44.3.681

Lowenstein, J. A., Blank, H., & Sauer, J. (2010). Uniforms affect the accuracy of children's eyewitness identification decisions. *Journal of Investigative Psychology and Offender Profiling*, *7*, 59–73. doi:10.1002/jip.104

Malloy, L. C., & Quas, J. A. (2009). Children's suggestibility: Areas of consensus and controversy. In K. Kuehnle & M. Connell (Eds.), *The evaluation of child sexual abuse allegations: A comprehensive guide to assessment and testimony* (pp. 267–297). Hoboken, NJ: Wiley.

Morgan, C. A., Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., … Southwick, S. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *International Journal of Law & Psychiatry*, *27*, 265–279. doi:10.1016/j.ijlp.2004.03.004

Nathanson, R., & Saywitz, K. J. (2003). The effects of courtroom context on children's memory and anxiety. *Journal of Psychiatry and Law*, *31*, 67–98.

Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: A social cultural developmental theory. *Psychological Review*, *111*, 486–511. doi:10.1037/0033-295X.111.2.486

Parker, J. F., & Carranza, L. E. (1989). Eyewitness testimony of children in target-present and target-absent lineups. *Law and Human Behavior*, *13*, 133–149. doi:10.1007/BF01055920

Parker, J. F., & Ryan, V. (1993). An attempt to reduce guessing behavior in children's and adults' eyewitness identification. *Law and Human Behavior*, *17*, 11–26. doi:10.1007/BF01044534

Peters, D. P. (1991). The influence of stress and arousal on the child witness. In: J. Doris (Ed.), *The suggestibility of children's recollections* (pp. 60–91). Washington, DC: American Psychological Association.

Pozzulo, J. D., & Lindsay, R. C. L. (1997). Increasing correct identifications by children. *Expert Evidence*, *5*, 126–132. doi:10.1023/A:1008875802767

Pozzulo, J. D., & Lindsay, R. C. L. (1998). Identification accuracy of children versus adults: A meta-analysis. *Law and Human Behavior*, *22*, 549–570. doi:10.1023/A:1025739514042

Quas, J. A., Bauer, A., & Boyce, T. (2004). Physiological reactivity, social support, and memory in early childhood. *Child Development*, *75*, 797–814. doi:10.1111/j.1467-8624.2004.00707.x

Quas, J. A., Carrick, N., Alkon, A., & Boyce, W. T. (2006). Children's memory for a mild stressor: The role of sympathetic activation and parasympathetic withdrawal. *Developmental Psychobiology*, *48*, 686–702. doi:10.1002/dev.20184

Quas, J. A., & Klemfuss, J. Z. (in press). Physiological stress reactivity and episodic memory in children. In P. Bauer & R. Fivush (Eds.), *Handbook of memory development*. Oxford, UK: Oxford University Press.

Quas, J. A., & Lench, H. C. (2007). Arousal at encoding, arousal at retrieval, interviewer support, and children's memory for a mild stressor. *Applied Cognitive Psychology*, *21*, 289–305. doi:10.1002/acp.1279

Quas, J. A., Rush, E. B., Yim, I., & Nikolayev, M. (2013). Stress, social support, and children's memory: Testing causal connections. *Memory*. Advance online publication. doi:10.1080/09658211.2013.809766

Quas, J. A., Yim, I. S., Rush, E., & Sumaroka, M. (2012). Hypothalamic pituitary adrenal axis and sympathetic activation: Joint predictors of memory in children, adolescents, and adults. *Biological Psychology*, *89*, 335–341. doi:10.1016/j.biopsycho.2011.11.006

Roozendaal, B. (2002). Stress and memory: Opposing effects of glucocorticoids on memory consolidation and memory retrieval. *Neurobiology of Learning & Memory*, *78*, 578–595. doi:10.1006/nlme.2002.4080

Saywitz, K. J., & Nathanson, R. (1993). Children's testimony and their perceptions of stress in and out of the courtroom. *Child Abuse and Neglect*, *17*, 613–622. doi:10.1016/0145-2134(93)90083-H

Scherer, D. G. (1991). The capacities of minors to exercise voluntariness in medical treatment decisions. *Law and Human Behavior*, *15*, 431–449. doi:10.1007/BF02074080

Schwabe, L., Joëls, M., Roozendaal, B., Wolf, O. T., & Oitzl, M. S. (2011). Stress effects on memory: An update and integration. *Neuroscience & Biobehavioral Reviews*, *36*, 1740–1749. doi:10.1016/j.neubiorev.2011.07.002

Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, *100*, 139–156. doi:10.1037/0033-2909.100.2.139

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior*, *22*, 603–647. doi:10.1023/A:1025750605807

Wells, G. L., & Turtle, J. W. (1986). Eyewitness identification: The importance of lineup models. *Psychological Bulletin*, *99*, 320–329. doi:10.1037/0033-2909.99.3.320

Wise, R. A., Safter, M. A., & Maro, C. M. (2011). What U.S. law enforcement officers know and believe about eyewitness factors, eyewitness interviews and identification procedures. *Applied Cognitive Psychology*, *25*, 488–500. doi:10.1002/acp.1717

Wogalter, M. S., Malpass, R. S., & McQuiston, D. W. (2004). A national survey of U.S. police on preparation and conduct of identification lineups. *Psychology, Crime & Law*, *10*, 69–82. doi:10.1080/10683160410001641873

Wolf, O. T. (2009). Stress and memory in humans: Twelve years of progress? *Brain Research*, *1293*, 142–154. doi:10.1016/j.brainres.2009.04.013

Wright, D. B., London, K., & Waechter, M. (2010). Social anxiety moderates memory conformity in adolescents. *Applied Cognitive Psychology*, *24*, 1034–1045. doi:10.1002/acp.1604

Yim, I. S., Quas, J. A., Cahill, L., & Hayakawa, C. M. (2010). Children's and adults' salivary cortisol responses to an identical psychosocial laboratory stressor. *Psychoneuroendocrinology*, *35*, 241–248. doi:10.1016/j.psyneuen.2009.06.014