

## Adults' ability to detect deception of stressful and non-stressful stories of children

Christine Saykaly<sup>a\*</sup>, Victoria Talwar<sup>a</sup>, Rod C.L. Lindsay<sup>b</sup>, Nicholas C. Bala<sup>c</sup>,  
Kang Lee<sup>d</sup>, Michelle Bertrand<sup>e</sup> and Sarah Michelle Nugent<sup>b</sup>

<sup>a</sup>*Department of Educational and Counselling Psychology, McGill University, Montreal, Quebec, Canada;* <sup>b</sup>*Psychology, Queen's University, Kingston, Ontario, Canada;* <sup>c</sup>*Faculty of Law, Queen's University, Kingston, Ontario, Canada;* <sup>d</sup>*OISE, University of Toronto, Toronto, Ontario, Canada;* <sup>e</sup>*Criminal Justice Department, University of Winnipeg, Winnipeg, Manitoba*

(Received 27 July 2011; final version received 13 April 2012)

This study examined adults' ability to detect deception of children who provided true and fabricated descriptions of stressful and non-stressful events. Observers viewed video clips of children telling four different stories (true non-stressful, true stressful, false non-stressful, false stressful). The story conditions were presented by either four different children (Study 1), or by the same child (Study 2). Results indicate that overall accuracy of discriminating non-stressful events is greater when viewing stories given by the same child (above chance levels) compared to four different children (at chance levels). However no differences occur for stressful events. The current study supports the need for ecologically valid research examining the content of children's reports.

**Keywords:** child witness testimony; lie detection

### Introduction

In recent decades, North America has experienced an increase in children testifying in court proceedings, particularly in cases where there are allegations of physical or sexual abuse (Bala, Lee, Lindsay, & Talwar, 2000; Bruck, Ceci, & Hembrooke, 1998; Gray, 1993; Honts, 1994). Traditionally, children were viewed as unreliable or susceptible to the influence of others (Cashmore & Bussey, 1996). However, more recent research demonstrates that children are able to give highly accurate accounts and be competent witnesses in courtroom settings (Ceci & Bruck, 1995; Quas, Goodman, Ghetti, & Redlich, 2000). Nevertheless, there are a number of issues pertinent to the justice system that remain to be a concern such as child witness veracity (i.e., their honesty – whether they deliberately fabricated or concealed evidence) and credibility (whether they are believed) (London & Nunez, 2002; Talwar, Lee, Bala, & Lindsay, 2004). These issues are of particular importance as there is a heightened concern that children may be coerced or coached to give false reports (Bala & Schuman, 1999; Brennan, 1994; Haugaard & Reppucci, 1992; Jones & McGraw, 1987).

Research has found that children can be coached to report fabricated events or falsely deny events that they actually experienced (Lyon, Malloy, Quas, & Talwar,

---

\*Corresponding author. Email: [christine.saykaly@mail.mcgill.ca](mailto:christine.saykaly@mail.mcgill.ca)

2008; Orcutt, Goodman, Tobey, Batterman-Faunce, & Thomas, 2001; Talwar, Lee, Bala, & Lindsay, 2006). Although researchers have found mixed results concerning adult's abilities to detect deceptive statements of children (e.g., Edelstein, Lutten, Ekman, & Goodman, 2006; Leach et al., 2009), few researchers have directly examined adults' abilities to assess the veracity of children's reports giving multi-statement *fabricated* reports (i.e. giving more than just a yes or no answer) (Lyon et al., 2008). Of the research that has been able to identify true from fabricated reports of children given elaborate responses, many have focused on the analysis of written transcripts (e.g. Bruck, Ceci, & Hembrook, 2002). Although this has significant implication, jurors are unable to analyze transcripts when making their decisions about the veracity of children's statements during court proceedings.

Whether adults are able to accurately assess the veracity of children's statements in court is an important issue to consider since in some court cases children's accounts may be particularly influential, as the final verdict may rest largely on whether a child is believed (Edelstein et al., 2006; Goodman, Golding, Helegen, Haith, & Michelli, 1987). In trials, jury members or judges must assess the veracity of a child's statements based on their perception and belief of the children's credibility (or lack thereof). Therefore, when considering the credibility of children's testimony, it is important to consider not only the children's abilities to testify in court, but also the ability of adults to distinguish between the true and false statements of children.

### ***Detection of children's true and fabricated reports***

Recent studies have found that children as young as three years are able to lie to deceive another, and can engage in other types of deceptive behaviour such as keeping a secret from others (Bottoms, Goodman, Schwartz-Kenney, & Thomas, 2002; Lyon et al., 2008; Talwar et al., 2004) or lying to protect themselves from getting into trouble (Lewis, Stranger, & Sullivan, 1989; Polak & Harris, 1999; Talwar & Lee, 2002, 2008). The majority of research on detection of children's deception has focused on the abilities of jurors, legal professionals, and especially lay-persons to 'intuitively' identify whether or not a child is being truthful or deceitful. Research on adult deception detection has found that most adults are unable to identify deceit at above chance levels (e.g., Bond & DePaulo, 2006; Vrij, 2000). However, some studies have found that observers may have greater facility at detecting deception in younger children (Chahal & Cassidy, 1995; Feldman & White, 1980; Westcott, Davies, & Clifford, 1991). For example, in a study by Feldman and White (1980), untrained observers were more accurate at judging the veracity of the statements of first graders compared to seventh graders and college students. Moreover, Westcott et al. (1991) found that when adult raters watched children's true and fabricated reports about a school trip, they were above chance levels at detecting deceptive statements of 7- and 8-year-old children. On the other hand, some studies report that adults are poor at detecting children's false reports (Lewis et al., 1989; Strömwall & Granhag, 2005; Talwar & Lee, 2002; Talwar et al., 2006). For instance, Talwar et al. (2006) found that adults' detection of children's true and false statements was at chance level. However, their perceptions of children's veracity varied with the age of the child.

A significant proportion of deception detection studies have asked participants to judge the veracity of children making one or two short statements (Leach, Talwar, Lee, Bala, & Lindsay, 2004; Lewis et al., 1989; Talwar & Lee, 2002). While these

studies are similar to everyday lies, they fail to replicate the longer reports given by children during court proceedings and forensic interviews. Furthermore, adults may be sensitive to differences between children's true and false reports when narratives are longer and involve more detail (Craig, Scheibe, Raskin, Kircher, & Dodd, 1999; Goodman & Reed, 1986; Pipe, Lamb, Orbach, & Esplin, 2004; Vrij, 2005). For example, in one study, Orcutt et al. (2001) found that mock jurors who saw children (aged 7–9 years) give detailed 'testimony' regarding a fabricated event (about a sticker being placed on their skin versus on top of their shirt) were more willing to 'convict' a male defendant on the basis of children's accurate testimony of the defendant's guilty actions than children who gave false testimony about the alleged event.

Other studies examining detection rates of older children's true and fabricated reports have yielded deception detection rates at or near chance levels (Crossman & Lewis, 2006; Leippe, Manion, & Romanczyk, 1992; Strömwall & Granhag, 2005; Strömwall, Granhag, & Landström, 2007; Vrij, Akehurst, Brown, & Mann, 2006). It is important to consider the context of the stories being reported and to judge the veracity of children's statements who are reporting about issues that are forensically relevant.

### *Children's experience of stressful events*

The few studies that examined the veracity of children's testimony or the ability of people to discriminate between 'real' and fabricated reports have generally focused on reports of mundane everyday events. However, when children's veracity is of interest to the judicial system, the situation is much more serious and the events are usually stressful or emotional to discuss. Fabrication of such reports may be more difficult for witnesses to maintain during questioning and therefore their statements would be easier to detect, even among older children. Previous studies that have examined children's reports of innocuous everyday events have found younger children's fabricated reports easier to detect than older children's (e.g., Bottoms et al., 2002; Goodman et al., 2006; Talwar & Lee, 2002). For example, Strömwall and Granhag (2005) asked children to give true and false reports about attending a magic show. Little research on children's ability to fabricate convincing reports of stressful events has been empirically examined.

Related research examining children's reports of fear-provoking events reveal differences between children's truthful reports of fear-provoking and non-fear-provoking events (Fivush, Sales, Goldberg, Bahrack, & Parker, 2004; Hudson, Gebelt, Haviland, & Bentivegna, 1992). Furthermore, some believe that high stress situations are clearly impressed upon children's minds, resulting in particularly salient and accurate memory reports (Eisen, Goodman, Qin, Davis, & Crayton, 2007; Fivush et al., 2004; Goodman, Bottoms, Schwartz-Kenney, & Rudy, 1991; see Pipe et al., 2004, for a review), whereas others believe the stress of an event might impede one's ability to accurately recall the events that transpired (Peterson & Bell, 1996).

In Study 1, adult observers viewed video segments of four children reporting one of four different types of events: true-non-stressful events, fabricated-non-stressful events, true-stressful events, and fabricated-stressful events. In Study 2, observers were asked to rate the same four accounts provided by a single child, and determine whether each child was being truthful or not. Observers were also asked follow-up

questions to determine their opinion of the credibility and veracity of the child's statements. Stimuli were created resembling testimony of stressful and non-stressful events that children may discuss in court settings. Children told stressful stories about stressful events (e.g., being bullied) and non-stressful stories about non-stressful events (e.g., going camping).

As shown in existing research, there are differences between reports of stressful and non-stressful events (Hudson et al., 1992). Thus it was hypothesized that adult observers will discriminate truthful and false stories more accurately when stressful events are reported compared to non-stressful events. Second, based upon both the developmental literature which has found that younger children are less sophisticated liars (e.g., Talwar & Lee, 2002) and studies that have reported better adult detection of younger children's lies (e.g., Bottoms et al., 2002; Goodman et al., 2006; Talwar & Lee, 2002), it was hypothesized that the adults' accuracy rates at discriminating between the true and fabricated reports seen in the videos would be higher for younger than older children.

## Study 1

### *Method*

#### *Participants*

Adult participants (herein referred to as observers) were recruited through an undergraduate participant pool. All 48 adult observers were undergraduate students between the ages of 18 and 28 years ( $M = 21.64$  years,  $SD = 2.08$ ). All were compensated for their participation by receiving an extra percent on their psychology course final grade. Seventy-three percent of the sample was female.

#### *Materials*

Experimental stimuli consisted of videotaped interviews of 36 children between the ages 4 and 9 years ( $M = 6.81$ ,  $SD = 1.88$ , 23 females) discussing both true and fabricated stressful and non-stressful events. Children were divided into three age groups: 4- or 5-year-olds ( $n = 12$ ), 6- or 7-year-olds ( $n = 12$ ), and 8- or 9-year-olds ( $n = 12$ ). Children were recruited from a database of children who had participated in previous studies and whose parents indicated willingness to participate in future studies. The majority of children were from white middle- and upper middle-income families in a large North American city. All child participants were brought to a university research laboratory by their parents.

Children were asked to provide four stories: a true non-stressful event, a false non-stressful event, a true stressful event, and a false stressful event. Parents were asked to identify various stressful (e.g., being bullied, a visit to the emergency room, having lost a parent in a crowded public area) and non-stressful (e.g., a camping trip, a birthday party, a family outing to the zoo) events that their children had or had not experienced. Children were asked whether they felt the stories were stressful or non-stressful in order to validate the selection of their parents. In the true condition, children were told to discuss an event that they attended or participated in and could remember. In the fabricated condition, children were asked to discuss an event in which they had never participated. True and false conditions were yoked in order for

the false events to be similar in 'stress' level as reported by the children. Each child told all four types of stories.

Parents assisted their children in creating or rehearsing their true and fabricated stories in preparation for their participation (for approximately 15 minutes). Children were instructed that they should try to be as convincing as possible during the true and fabricated story-telling, as the interviewer was blind to which stories were true and which were fabricated. Children were interviewed about the 'alleged event' by a research assistant, who prompted children with open-ended questions (e.g., 'Can you tell me about the trip you made to the zoo?' 'How did you feel when [event] happened?'). Questions remained the same for all conditions. The stories were edited in order to eliminate any identifying information. Video clips of the interviews varied in length from two to six minutes ( $M = 3.31$ ,  $SD = 1.16$ ). There were no significant differences in narrative length between the four types of stories. Preliminary analyses revealed no relation between adult accuracy and story length ( $p > .05$ ) for any of the four story types.

### *Procedure*

Adult observers were tested individually in a quiet room, with sessions lasting approximately 30 minutes. A 2 (Veracity: true or false)  $\times$  2 (Story Type: non-stressful or stressful)  $\times$  3 (Age of Child: 4- or 5-year-olds, 6- or 7-year-olds, 8- or 9-year-olds) repeated measures design (where age of child as the only between subject variable) was used to examine the observers' detection accuracy of children's reports. Using MediaLab software, each observer saw four video clips, one per story type condition (true-non-stressful, true-stressful, fabricated-non-stressful, fabricated-stressful) given by four different children. The four conditions were viewed in random order. Observers were instructed that they would be watching videos of children reporting about a past event and that the stories may be true or untrue. Prior to viewing the videos, observers were asked to complete a demographic questionnaire.

After viewing each video clip, observers were asked several questions about the credibility and veracity of the child they just viewed. To examine adults' abilities when assessing the veracity of the reports they were asked: 'Overall (ignoring minor errors or omissions of detail and focusing only on the central claims of the story), do you believe that the events described by the child actually happened?' Observers had to respond 'yes' or 'no'. Observers were asked to rate their certainty in their decision regarding whether they believed that the event described by that child actually took place on a 10-point scale (1 = *not at all certain* to 10 = *absolutely certain*). Observers were asked to rate to what degree they believed the described event would have been a stressful or non-stressful experience for the child using a 10-point scale (1 = *very non-stressful* to 10 = *very stressful*). This question served as a manipulation check that examined whether participants accurately perceived the events to be stressful or non-stressful.

To determine whether there were discernible, nonverbal expressive behaviour differences among the target children, children's nonverbal expressive behaviours were coded by research assistants. In general, children did not display much expressive behaviour, similar to previous research (e.g., Lewis et al., 1989; Talwar & Lee, 2002). Therefore, some of the more frequently occurring movements were collapsed into two categories: positive movements and negative movements (Talwar, Gordon, & Lee,

2007). Positive movements included seven categories associated with vertical musculature movements that elongated the face in some way (e.g., lip-corner pulling, brow raises). Negative movements included nine categories associated with vertical movements that scrunched the face in some way (e.g., lip tightening, brow furrowing).

### Results

Preliminary analyses yielded no significant differences in terms of the sex of the observers. Thus, the results from both sexes were combined for all subsequent analyses. A series of 2 (Veracity)  $\times$  2 (Story Type)  $\times$  3 (Age of Child) analyses of variances (ANOVAs) were performed for each of the follow-up questions asked immediately after each clip was viewed in order to determine observer's discrimination abilities and perceived beliefs regarding each child's report.

#### Accuracy

Observers were asked after each video clip whether they believed the events described by the child actually happened. Observer accuracy was calculated based upon whether the child was indeed telling a true or fabricated story. Adult observers could make a correct decision in two ways: they believed the children when they were telling the truth, or did not believe the children when they were telling a false story. Observers' overall accuracy and their accuracy of the four story conditions were calculated. Overall accuracy across all four categories was found to be 56.2%, which was not significantly above chance (see Table 1). One-sample *t*-tests were used to compare accuracy to the level of chance (50%) for each story condition. As seen in Table 1, observer accuracy for true non-stressful stories (79.0%) was significantly above chance level while accuracy for true stressful stories (35.4%) was significantly below chance level. Accuracy for false non-stressful stories (62.5%) and false stressful stories (41.7%) were not significantly different from chance. Accuracy judgements were then entered in a 2 (Veracity)  $\times$  2 (Story Type)  $\times$  3 (Age of Child) ANOVA which yielded a significant main effect for Story Type,  $F(1, 45) = 22.63$ ,  $p < .001$ ,  $\eta^2 = .335$ . Observers had a higher accuracy rate for non-stressful stories (70.8%) compared to stressful stories (38.5%). There were no other main effects and no significant interactions.

Table 1. Discrimination accuracy.

Accuracy	Study 1				Study 2			
	<i>M</i>	<i>SD</i>	<i>t</i> (47)	Sig	<i>M</i>	<i>SD</i>	<i>t</i> (71)	Sig
Overall	.563	.48	1.47	.147	.587	.30	2.45	.017*
True Stressful	.354	.50	-2.09	.042*	.542	.50	.705	.483
False Stressful	.417	.41	-1.16	.252	.444	.50	-.942	.349
True Non-stressful	.790	.49	4.92	.000***	.681	.47	3.263	.002**
False Non-stressful	.625	.29	1.77	.083	.583	.47	1.42	.159

\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ .

### *Confidence of assessments*

Observers were asked how confident they were in their assessment of whether the events being described by the child actually happened. No significant differences were found. Overall, approximately 81.2% felt somewhat to moderately confident in their judgments.

### *Perceived stressfulness and non-stressfulness*

Adults rated how stressful or non-stressful they believed the described events were on a scale of 1 (non-stressful) to 10 (stressful). The ratings were entered into a 2 (Veracity)  $\times$  2 (Story Type)  $\times$  3 (Age of Child) ANOVA. A significant main effect was found for Story Type,  $F(1, 45) = 345.60$ ,  $p < .001$ ,  $\eta_p^2 = .885$ . Observers rated the stressful stories as significantly more stressful ( $M = 7.02$ ,  $SE = .21$ ) than the non-stressful stories ( $M = 2.15$ ,  $SE = .13$ ). There were no other significant main effects. However, there was a significant Story Type  $\times$  Age of Child interaction  $F(2, 45) = 4.28$ ,  $p \leq .05$ ,  $\eta_p^2 = .16$ . The 8- or 9-year-old children's non-stressful reports ( $M = 1.44$ ,  $SE = .22$ ) were rated as least stressful in comparison to reports by children in the 4- or 5-year-old ( $M = 2.84$ ,  $SE = .22$ ) and the 6- or 7-year-old groups ( $M = 2.16$ ,  $SE = .22$ ). Both the 6- or 7-year-old ( $M = 7.38$ ,  $SE = .37$ ) and 8- or 9-year-old children's stressful reports ( $M = 7.03$ ,  $SE = .37$ ) were also rated as more stressful compared to 4- or 5-year-old reports ( $M = 6.66$ ,  $SE = .37$ ).

### *Non-verbal coding*

Analyses revealed significant differences among children's stories for negative expressive behaviour,  $F(1, 35) = 5.26$ ,  $p < .05$ . There was a difference between children's true ( $M = .86$ ,  $SD = .98$ ) and fabricated stressful stories ( $M = 3.47$ ,  $SD = 2.18$ ) ( $p < .05$ ) while children's true ( $M = 1.21$ ,  $SD = 1.19$ ) and fabricated non-stressful stories ( $M = .94$ ,  $SD = 1.13$ ) were similar. There were no significant differences in positive expressive behaviour (true non-stressful:  $M = 1.24$ ,  $SD = 1.15$ ; false non-stressful:  $M = 1.08$ ,  $SD = 1.17$ ; true stressful:  $M = 1.63$ ,  $SD = 1.23$ ; false stressful:  $M = 1.94$ ,  $SD = 1.43$ ). Thus, children's expressive behaviours did differ in terms of their negative expressive behaviour when giving true and false stressful accounts.

### *Discussion*

Overall, observers made accurate judgements for only 56.25% of the videos, which was not significantly above chance. This finding is similar to previous studies indicating that adult deception detection abilities are at chance levels (e.g., Bond & DePaulo, 2006; Vrij, 2000). It was hypothesized that observers would discriminate between true and fabricated stories of stressful events more accurately than non-stressful events. The findings of the current study were contrary to the stated hypothesis, as observers discriminated non-stressful events with greater accuracy than stressful events. Observers were able to correctly identify true non-stressful stories significantly above chance levels, and were significantly below chance levels in their accurate detection of true stressful events. That is, when children described

stressful events, observers appeared to think that these events were unlikely to have happened when the reports were actually true. Detection of fabricated stressful and fabricated non-stressful events did not vary from chance. Therefore, observers tended to show a bias towards believing children's reports of non-stressful events as compared to stressful events. This has important implications as children who testify during court proceedings almost always report on stressful events (such as abuse and neglect) (Bala et al., 2000; Bruck et al., 1998). Furthermore, if lay-people must evaluate the veracity of such stressful statements (as in the case of jurors), our findings suggest that they may be more sceptical about statements concerning stressful events and have a harder time assessing them. Importantly, lay-people also appear to have difficulty discriminating fabricated events, regardless of the content of their accounts, which leads one to question whether jurors in actual court proceedings would show similar detection tendencies.

It was hypothesized that adults would be able to discriminate the stories of younger children with greater success than older children. This hypothesis was not supported as no differences were found in detection rates with regards to the age of the child, which is consistent with other studies that also report that even young children's deceitful statements can be difficult to detect (e.g., Leach et al., 2004). However, older children's reports were perceived to be more stressful or non-stressful, respectively. This suggests that there were differences in the quality of children's reports based on age. However, adults were unable to use these differences to aid them in the detection of children's true and fabricated reports. Although unable to accurately discriminate between children's true and fabricated reports, adults were moderately confident in their ratings about their decision of the veracity of children's statements. Similar to previous findings (Leach et al., 2004), these results suggest that observer's perceived confidence in their abilities to detect deception is not reflected in their actual abilities.

Analyses of non-verbal behaviour suggest that the children did display some subtle cues to their deception. However, it appears that adults were not able to perceive or interpret these cues. Children did display differences in their true and fabricated stressful stories with more negative expressive behaviours demonstrated in their fabricated stressful stories. However, adults tended to be the least accurate at detecting children's true stressful stories. This may reflect a bias not to believe children's stressful stories. Adults may have expected children to show more negative expressive behaviours when describing the stressful events. On the other hand, adults were most accurate at detecting children's truthful non-stressful stories. However, children did not display significant expressive behaviours in their true non-stressful stories that distinguished them from other stories. This suggests that adults' decisions were not based on their attention to expressive cues. Instead, the results suggest that adults appeared to rely on biases to decide which stories to believe. Adults may be biased to believe that children are more likely to tell the truth about non-stressful events, possibly because they have more day-to-day familiarity with children's true non-stressful stories compared to the other types of stories viewed.

Observers of the current study saw video clips of four different children. It is possible that children's true and fabricated reports would be easier to detect had adults viewed clips of the same child discussing all four event types. It has been suggested that raters observing within-subjects stimuli may have better discrimination of true and fabricated reports. Police often claim that they see witnesses more

than once in various contexts, and therefore get a 'feel' for when such an individual is being truthful or deceitful. By seeing the witness make multiple statements, they develop familiarity with the individual and a heightened awareness for the individual's behaviours which can aid in their judgments regarding the veracity of their statements. It follows then that when faced with multiple statements given by the same individual, deception detection may be improved. Some evidence for this comes from a study by DePaulo, Stone, and Lassiter (1985) where adult participants were asked to give two truthful statements and two false statements about controversial issues (e.g., abortion). Observers viewed the four accounts of an individual and were asked to rate the sincerity of the arguments. Results indicated that participants perceived the false statements as less sincere than the truthful ones. Thus, it may be that seeing one child's four stories may increase adult raters' detection abilities of the children's true and false statements. Therefore, in Study 2, adults' detection abilities when they viewed all four conditions of the same child were examined.

## Study 2

### *Method*

#### *Participants*

Observers were 72 students (undergraduate students) recruited from an undergraduate participant pool. Observers ranged in age from 15 to 45 years ( $M = 19.39$ ,  $SD = 4.51$ ). All were compensated for their participation by receiving an extra percent on their psychology course final grade or by receiving \$10 (CND). Approximately 75% of the sample was female.

#### *Materials*

The stimuli for this experiment were the same as Study 1. Observers viewed four clips, all of the same child reporting on all four conditions. Two different adult observers viewed each child's story clips. The clips from Study 1 and Study 2 were the same, simply regrouped for Study 2. Conditions were presented in a random order.

#### *Procedure*

Observers completed a demographic questionnaire. With the use of MediaLab software, observers viewed four video clips, one per story condition, and were asked the same follow-up questions as Study 1. Unlike Study 1 where observers viewed clips of four different children, observers in Study 2 viewed the same child given in all four story clips. Each session took approximately 30 minutes to complete. Observers were unable to change their responses upon seeing additional clips of the child.

#### *Design*

Study 2 was a 2 (Veracity: true or false)  $\times$  2 (Story Types: non-stressful or stressful)  $\times$  3 (Age of Child: 4- or 5-year-olds, 6- or 7-year-olds, 8- or 9-year-olds) repeated measures design (with age as the only between-subject variable).

## Results

### Accuracy

Similar to Study 1, observers were asked after each video clip whether they believed the events described by the child actually happened. Overall accuracy across all four categories was found to be 58.7%, a rate which was significantly above chance (see Table 1). One-sample *t*-tests were used to compare accuracy to the level of chance (50%) for each story condition. As seen in Table 1, observer accuracy for true non-stressful stories (68.1%) was significantly above chance levels. Accuracy for true stressful stories (54.2%), fabricated non-stressful stories (58.3%), and fabricated stressful stories (44.4%) was not significantly different from chance. Accuracy judgements were then entered in a 2 (Veracity)  $\times$  2 (Story Type)  $\times$  3 (Age of Child) ANOVA which yielded a significant main effect of Type of Story,  $F(1, 69) = 5.29, p < .05, \eta_p^2 = .071$ . Observers had a higher accuracy rate for non-stressful stories (63.2%) compared to stressful stories (49.3%). No other significant differences were found.

To examine whether accuracy improved when seeing all four video clips of a single child, analyses were conducted to examine if accuracy improved between seeing the first video clip and subsequent video clips regardless of story or veracity condition. There was a significant difference between seeing the first video clip and the final video clip,  $t(71) = 2.76, p = 0.007$ . There were no significant differences between the first video clip and the second or third video clips ( $ps > .05$ ) or accuracy between subsequent interviews (e.g., between interview 2 and 3, 2 and 4, or 3 and 4;  $ps > .05$ ). Thus, it appears that accuracy improved overall between seeing the first video clip ( $M = .44, SD = .50$ ) and the fourth video clip ( $M = .68, SD = .47$ ). This difference was not seen in Study 1. Furthermore, Study 2's accuracy for the first interview and Study 1's accuracy for the first interview were compared and there was no significant difference.

### Confidence of assessment

Observers were asked how confident they were in their assessments of whether the events being described by the child actually happened. No significant differences were found. Overall, approximately 86.6% of adults felt moderately confident to confident in their judgments.

### Perceived stressfulness and non-stressfulness

The story event manipulation check (i.e., the degree to which the event would have been a non-stressful or stressful experience for the child) was examined. The ratings were entered into a 2 (Veracity)  $\times$  2 (Story Type)  $\times$  3 (Age of Child) ANOVA. Observers rated the stressful stories as significantly more stressful ( $M = 6.62, SE = .17$ ) than the non-stressful stories ( $M = 1.71, SE = .16$ ),  $F(1, 69) = 342.45, p < .001, \eta_p^2 = .83$ . No other significant main effects or interactions were found.

## Discussion

Similar to the results of Study 1, observers were more accurate at discriminating the veracity of non-stressful stories compared to stressful stories. Furthermore, observers

were significantly above chance when asked to judge children's true non-stressful events. Notably, in Study 2, participants' overall deception detection accuracy reached levels significantly above chance. Most importantly, results indicated that viewing the same child more than once helped adults discriminate between children's true and fabricated reports. Specifically, the difference between the first and fourth viewing was significant. It is possible that viewing repeated clips of one child might have increased the adults' familiarity with the child as well as revealing cues to the child's deception. Such findings would confirm the suggestions by police officers and other front-line workers that seeing multiple statements of a witness can increase abilities to judge veracity. Therefore, although the difference in overall accuracy between Study 1 (56.2%) and Study 2 (58.7%) was not significant, based on the improvement in accuracy as participants viewed successive clips of the same child, there may be improvements in the overall accuracy in deception detection when viewing multiple statements of the same child.

Furthermore, similar to Study 1, no differences were found based on age. Adults were not able to discriminate the veracity of younger children's reports more easily than that of older children. However, as in Study 1, adults were again confident in their abilities to discriminate true and fabricated reports, despite not being accurate in their ratings.

### General discussion

The present set of experiments examined whether adults could detect children's true and fabricated reports about non-stressful and stressful events. Results from both studies suggest that adults were better at correctly discriminating the veracity of children's reports of non-stressful events than stressful events. Children's true non-stressful stories were detected significantly better than expected by chance in both studies whereas adults were less accurate at detecting children's true stressful stories with performance below chance discrimination levels in Study 1. This finding is noteworthy as it suggests that adults might be unable to accurately detect children's reports of true stressful events, which are the types of events children most commonly testify about in court. Adults were significantly more accurate at detecting children's reports of non-stressful events, which are the types of events children rarely report in court. These findings raise concerns about detection accuracy by jurors in legal settings as children are typically not giving reports about mundane everyday events but rather reports of stressful events.

There are two important differences between the current methodology and previous work on deception detection of children. First, the length of the narratives

Table 2. Perceived stressfulness and non-stressfulness of statements.

Age	Non-stressful		Stressful	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Preschool	2.84	.218	6.66	.365
Early Elementary	2.16	.218	7.38	.365
Older Elementary	1.44	.218	7.03	.365

seen by observers was greater than in previous work. Specifically, observers were asked to view children ‘testifying’ for two to six minutes, whereas the stimuli used in previous studies often were shorter statements (e.g., Leach et al., 2004; Talwar & Lee, 2002). Also, the current stimuli more closely resemble the reports of children being questioned about a previous event. Second, the current stimuli were generated by asking children to fabricate stories while previous studies were based on an assessment of the veracity of children ‘spontaneously’ lying about events in a lab setting (e.g., Crossman & Lewis, 2006; Talwar & Lee, 2008). At times, when children are not telling the truth in court, it is because they have been encouraged to lie by an adult (such as a parent). Children in these cases are encouraged to fabricate a story, which is different from a spontaneous lie (e.g., lying to get out of trouble because you peeked at a toy). Therefore, the stimuli of the current research are forensically relevant making this study more ecologically valid than some previous studies.

It is also important to note that although children reported stressful events, it is likely that these events did not elicit the same stress levels in children as events that they may report of in court, such as abuse, or in school, such as bullying. Within the ethical limitations that arise from such research, the current methodology attempted to examine children’s reports of stressful events that were analogous to the types of events they may be reported in legal cases. Furthermore, as these narratives may not have elicited as much stress as, for example, reporting about sexual abuse in court, it is possible that observers experienced difficulties in judging the veracity of statements due to the lack of emotional response by the children when telling their stressful stories; research has shown that observers have a tendency to believe that victims of abuse may display emotions such as crying (Regan & Baker, 1998). Nevertheless, further research is needed to determine whether true and false accounts of emotionally harmful events, even more similar to the types of testimony children give in court, would generate comparable discrimination results as were found in the current studies.

Although the current stimuli bore some resemblance to children testifying in court, there was no cross-examination. Understanding adult’s perceptions of children’s statements during direct and cross-examination may yield different effects. Future research might also benefit from looking at stressful and non-stressful accounts during both questioning styles. One must note that the current sample of observers were mostly female. Although this is representative of the population of psychology undergraduate students, the lack of male observers should be noted. Although unlikely, having a more balanced observer sample may yield different results.

Nevertheless, testifying in court can, by itself, be stress-inducing and may affect children’s abilities to recount events accurately. The stress children feel when testifying may interact with the type of event they are reporting such that differences between their true and fabricated accounts may be enhanced or diminished. Further research is needed to examine the effects of stress on children’s veracity in ecologically relevant settings and adults’ abilities to judge the credibility of children’s reports of stressful events. The current study suggested that adults have difficulty discriminating between children’s true and false reports when they are describing stressful events. Thus, child witnesses reporting events such as abuse or victimization may be less likely to be believed, even when their accounts are truthful.

## References

- Bala, N., & Schuman, J. (1999). Allegations of sexual abuse when parents have separated. *Canadian Family Law Quarterly*, 17, 191–241.
- Bala, N., Lindsay, R.C.L., Lee, K., & Talwar, V. (2000). A legal and psychological critique of the present approach to the assessment of the competence of childwitnesses. *Osgoode Hall Law Journal*, 38, 409–452.
- Bond, C.F., & DePaulo, B.M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, 10, 214–234. doi: [10.1207/s15327957pspr1003\\_2](https://doi.org/10.1207/s15327957pspr1003_2)
- Bottoms, B.L., Goodman, G.S., Schwartz-Kenney, B.M., & Thomas, S.N. (2002). Understanding children's use of secrecy in the context of eyewitness reports. *Law and Human Behavior*, 26, 285–313. doi: [10.1023/A:1015324304975](https://doi.org/10.1023/A:1015324304975)
- Brennan, M. (1994). The battle for credibility: Themes in the cross-examination of child victim witnesses. *International Journal of Semiotics of Law*, 7, 51–73. doi: [10.1007/BF01099010](https://doi.org/10.1007/BF01099010)
- Bruck, M., Ceci, S.J., & Hembrooke, H. (1998). Reliability and credibility of young children's reports. *American Psychologist*, 53, 136–151. doi: [10.1037/0003-066X.53.2.136](https://doi.org/10.1037/0003-066X.53.2.136)
- Bruck, M., Ceci, S.J., & Hembrooke, H. (2002). The nature of children's true and false narratives. *Developmental Review*, 22, 520–544. doi: [http://dx.doi.org/10.1016/S0273-2297\(02\)00006-0](http://dx.doi.org/10.1016/S0273-2297(02)00006-0).
- Cashmore, J., & Bussey, K. (1996). Judicial perceptions of child witness competence. *Law and Human Behaviour*, 20, 313–334. doi: [10.1007/BF01499026](https://doi.org/10.1007/BF01499026)
- Ceci, S.J., & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children's testimony*. Washington, DC: American Psychological Association.
- Chahal, K., & Cassidy, T. (1995). Deception and its detection in children: A study of adult accuracy, psychology. *Crime & Law*, 1, 237–245. doi: [10.1080/10683169508411959](https://doi.org/10.1080/10683169508411959)
- Craig, R.A., Scheibe, R., Raskin, D.C., Kircher, J.C., & Dodd, D. (1999). Interviewer questions and content analysis of children's statements of sexual abuse. *Applied Developmental Science*, 3, 77–85. doi: [10.1207/s1532480xads0302\\_2](https://doi.org/10.1207/s1532480xads0302_2)
- Crossman, A.M., & Lewis, M. (2006). Adults' ability to detect children's lying. *Behavioral Sciences and the Law*, 24, 703–715. doi: [10.1002/bsl.731](https://doi.org/10.1002/bsl.731)
- DePaulo, B.M., Stone, J.I., & Lassiter, G.D. (1985). Telling ingratiating lies: Effects of target sex and target attractiveness on verbal and nonverbal deceptive success. *Journal of Personality and Social Psychology*, 48, 1191–1203. doi: [10.1037/0022-3514.48.5.1191](https://doi.org/10.1037/0022-3514.48.5.1191)
- Edelstein, R.S., Luten, T.L., Ekman, P., & Goodman, G.S. (2006). Detecting lies in children and adults. *Law and Human Behavior*, 30, 1–10. doi: [10.1007/s10979-006-9031-2](https://doi.org/10.1007/s10979-006-9031-2)
- Eisen, M., Goodman, G.S., Qin, J.J., Davis, S., & Crayton, J. (2007). Maltreated children's memory: Accuracy, suggestibility, and psychopathology. *Developmental Psychology*, 43, 1275–1294. doi: [10.1037/0012-1649.43.6.1275](https://doi.org/10.1037/0012-1649.43.6.1275)
- Feldman, R.S., & White, J.B. (1980). Detecting deception in children. *Journal of Communication*, 30, 121–128.
- Fivush, R., Sales, J.M., Goldberg, A., Bahrick, L., & Parker, J. (2004). Weathering the storm: Children's long-term recall of Hurricane Andrew. *Memory*, 12, 104–118. doi: [10.1080/09658210244000397](https://doi.org/10.1080/09658210244000397)
- Goodman, G.S., & Reed, R.S. (1986). Age differences in eyewitness testimony. *Law and Human Behavior*, 10, 317–332.
- Goodman, G.S., Bottoms, B.L., Schwartz-Kenney, B., & Rudy, L. (1991). Children's memory for a stressful event: Improving children's reports. *Journal of Narrative and Life History*, 1, 69–99.
- Goodman, G.S., Golding, J.M., Helegen, V.S., Haith, M.M., & Michelli, J. (1987). When a child takes the stand. *Law and Human Behavior*, 11, 27–40.
- Goodman, G.S., Myers, J.E.B., Qin, J., Quas, J.A., Castelli, P., Redlich, A., et al. (2006). Hearsay versus children's testimony: Effects of truthful and deceptive statements on jurors' decisions. *Law and Human Behavior*, 30, 363–401. doi: [10.1007/s10979-006-9009-0](https://doi.org/10.1007/s10979-006-9009-0)
- Gray, E. (1993). *Unequal justice: The prosecution of child sexual abuse*. New York, NY: MacMillan.
- Haugaard, J.J., & Reppucci, N.D. (1992). Children and the truth. In S.J. Ceci, M.D. Leichtman, & M. Putnick (Eds.), *Cognitive and social factors in early deception* (pp. 29–45). Hillsdale, NJ: Erlbaum.

- Honts, C.R. (1994). Assessing children's credibility: Scientific and legal issues in 1994. *North Dakota Law Review*, 70, 879–899.
- Hudson, J., Gebelt, J., Haviland, J., & Bentivegna, C. (1992). Emotion and narrative structure in young children's personal accounts. *Journal of Narrative and Life History*, 2, 129–150.
- Jones, D.P., & McGraw, J.M. (1987). Reliable and fictitious accounts of sexual abuse to children. *Journal of Interpersonal Violence*, 2, 27–45.
- Leach, A.M., Lindsay, R.C.L., Koehla, R., Beaudry, J.L., Bala, N., Lee, K., & Talwar, V. (2009). The reliability of lie detection performance. *Law and Human Behaviour*, 33, 96–109. doi: 10.1007/s10979-008-9137-9
- Leach, A.M., Talwar, V., Lee, K., Bala, N., & Lindsay, R.C.L. (2004). Intuitive lie detection of children's deception by law enforcement officials and university students. *Law and Human Behavior*, 28, 661–685. doi: 10.1007/s10979-004-0793-0
- Leippe, M.R., Manion, A.P., & Romanczyk, A. (1992). Eyewitness persuasion: How and how well do fact finders judge the accuracy of adults' and children's memory reports? *Journal of Personality and Social Psychology*, 63, 181–197.
- Lewis, M., Stanger, C., & Sullivan, M.W. (1989). Deception in 3-year-olds. *Developmental Psychology*, 25, 439–443. doi: 10.1037/0012-1649.25.3.439
- London, K., & Nunez, N. (2002). Examining the efficacy of true/lie discussions in predicting and increasing the veracity of children's reports. *Journal of Experimental Child Psychology*, 83, 131–147. doi: 10.1016/S0022-0965(02)00119-4
- Lyon, T.D., Malloy, L.C., Quas, J.A., & Talwar, V. (2008). Coaching, truth induction, and young maltreated children's false allegations and false denials. *Child Development*, 79, 914–929. doi: 0009-3920/2008/7904-0008
- Orcutt, H.K., Goodman, G.S., Tobey, A.E., Batterman-Faunce, J.M., & Thomas, S. (2001). Detecting deception in children's testimony; Fact finders' abilities to reach the truth in open court and closed-circuit trials. *Law and Human Behavior*, 25, 339–372.
- Peterson, C., & Bell, M. (1996). Children's memory for traumatic injury. *Child Development*, 67, 3045–3070.
- Pipe, M.E., Lamb, M.E., Orbach, Y., & Esplin, P.W. (2004). Recent research on children's testimony about experienced and witnessed events. *Developmental Review*, 24, 440–468. doi: 10.1046/j.dr.2004.08.006.
- Polak, A., & Harris, P.L. (1999). Deception by young children following noncompliance. *Developmental Psychology*, 35, 561–568. doi: 10.1037/0012-1649.35.2.561
- Quas, J.A., Goodman, G.S., Ghetti, S., & Redlich, A.D. (2000). Questioning the child witness: What can we conclude from the research thus far? *Trauma, Violence, & Abuse*, 1, 223–249. doi: 10.1177/1524838000001003002
- Regan, P.C., & Baker, S.J. (1998). The impact of child witness demeanor on perceived credibility and trial outcome in sexual abuse cases. *Journal of Family Violence*, 13, 187–195. doi: 10.1023/A:1022845724226
- Strömwall, L.A., & Granhag, P.A. (2005). Children's repeated lies and truths: Effects on adults judgements and reality monitoring scores. *Psychiatry, Psychology and Law*, 12, 345–356.
- Strömwall, L.A., Granhag, P.A., & Landström, S. (2007). Children's prepared and unprepared lies: Can adults see through their strategies? *Applied Cognitive Psychology*, 27, 457–471. doi: 10.1002/acp.1288.
- Talwar, V., & Lee, K. (2002). Development of lying to conceal a transgression: Children's control of expressive behavior during verbal deception. *International Journal of Behavioral Development*, 26, 436–444. doi: 10.1080/01650250143000373
- Talwar, V., & Lee, K. (2008). Social and cognitive correlates of children's lying behavior. *Child Development*, 79, 866–881. doi: 10.1111/j.1467-8624.2008.01164.x
- Talwar, V., Gordon, H., & Lee, K. (2007). Lying in the elementary school: Verbal deception and its relation to second-order belief understanding. *Developmental Psychology*, 43, 804–810. doi: 10.1037/0012-1649.43.3.804
- Talwar, V., Lee, K., Bala, N., & Lindsay, R.C.L. (2004). Children's lie-telling to conceal a parent's transgressions: Legal implications. *Law and Human Behavior*, 28, 441–435. doi: 10.1023/B:LAHU.0000039333.51399.f6.
- Talwar, V., Lee, K., Bala, N., & Lindsay, R.C.L. (2006). Adult's judgment of children's coached reports. *Law and Human Behavior*, 30, 561–570. doi: 10.1007/s10979-006-90388

- Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice*. West Sussex: Wiley.
- Vrij, A. (2005). Criteria-Based Content Analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law*, *11*, 3–41. doi: [10.1037/1076-8971.11.1.3](https://doi.org/10.1037/1076-8971.11.1.3)
- Vrij, A., Akehurst, L., Brown, L., & Mann, S. (2006). Detecting lies in young children, adolescents and adults. *Applied Cognitive Psychology*, *20*, 1225–1237.
- Westcott, H., Davies, G., & Clifford, B.R. (1991). Adult's perceptions of children's videotaped truthful and deceptive statements. *Children and Society*, *5*, 123–135.