

Adults' Detection of Deception in Children: Effect of Coaching and Age for Children's True and Fabricated Reports of Injuries

Kelly L. Warren^{*}, Aishah Bakhtiar[†], Brent Mulrooney[†],
Graham Raynor[†], Elyse Dodd[†] and Carole Peterson[†]

A total of 1,074 undergraduates judged the truthfulness of children's interviews (from verbatim transcripts) about experiencing injuries serious enough to require hospital emergency room treatment. Ninety-six children (three age groups: 5–7, 8–10, and 11–14 years, 50% girls) were interviewed. At each age, 16 children told truthful accounts of actual injury experiences and 16 fabricated their reports, with half of each group coached by parents for the previous 4 days. Lies by 5- to 7-year-olds, whether coached or not, were detected at above-chance levels. In contrast, 8- to 10-year-olds' accounts that were coached, whether true or not, were more likely to be believed. For 11- to 14-year-olds, adults were less likely to accurately judge lies if they were coached. The believability of children aged 8 or above who were coached to lie is particularly disturbing in light of the finding that participants were more confident in the accuracy of their veracity decisions when judging coached reports. Copyright © 2015 John Wiley & Sons, Ltd.

In recent years a great deal of attention has been directed towards determining how well adults judge the veracity of children's statements. This is important, as any juror who serves on a case involving a child will have to make a judgement about the believability of the child. Jurors' failures to identify true reports could have damaging effects on children who have been victims of abuse or other crimes, and perpetrators are left unpunished. In contrast, jurors' failures to detect lies might promote unfair punishments for defendants who are falsely alleged to have committed a crime. Either outcome is unacceptable, making it important to determine whether adults can make such judgements, and whether specific factors influence a person's ability to make the appropriate decisions on veracity. An issue that remains largely unexplored is the influence on believability of coaching children about an event.

A recent meta-analysis by Bond and DePaulo (2006) assessed adults' ability to determine the veracity of information provided by other adults. Results indicated that those who made these veracity decisions averaged a 54% accuracy rate. People made more accurate decisions if they had previous exposure to the person lying, but they made more inaccurate decisions if the people lying were given a chance to prepare their lie. Interestingly, the analyses showed that as a group, people were better at judging

^{*}Correspondence to: Kelly Warren, Assistant Professor, Psychology Program, Grenfell Campus, Memorial University of Newfoundland, 20 University Drive, Corner Brook, NL, A2H 5G4, Canada. E-mail: kwarren@grenfell.mun.ca

[†]Memorial University of Newfoundland, St. John's, NL, Canada

when someone was telling the truth than when someone was telling a lie. In addition, the medium through which information was presented influenced the accuracy of decisions. More specifically, Bond and DePaulo (2006) noted that the ability to make a truth–lie decision decreased if participants were presented with information in a video format, but that there was no difference in veracity decisions for audiovisual-, audio-, or transcript-based decisions. A second meta-analysis showed little relationship between confidence and the accuracy of veracity decisions (DePaulo *et al.*, 1997).

To the best of our knowledge, no meta-analysis has been completed comparing studies that have assessed the ability to determine the veracity of children's statements. However, the findings with children are similarly bleak. A review of the literature indicates accuracy ratings ranging from a low of 46.1% judgement accuracy in a study assessing children's prepared lies (Stromwall, Granhag, & Landstrom, 2007) to a high of 72% judgement accuracy in a study where the adults were presented with a video showing the script of the event children were lying about (Tye *et al.*, 1999). As with adults, the accuracy of veracity judgements appears to be dependent on the circumstances surrounding both the lies children are telling and the judgement process itself.

Content of the Lie

Studies assessing veracity decisions have varied greatly with respect to the content of children's lies. For example, children who played a guessing game were given the opportunity to surreptitiously peek when the experimenter left the room (Crossman & Lewis, 2006; Lewis, Stanger, & Sullivan, 1989; London & Nunez, 2002; Polak & Harris, 1999). The majority of children peeked, and many spontaneously lied about it when questioned. When video clips of the children lying or telling the truth about peeking were shown to adults who were asked to judge child truthfulness (Leach *et al.*, 2009; Talwar, Gordon, & Lee, 2007a; Talwar & Lee, 2002), observers' judgements were mostly quite poor. Other researchers have studied children's lies about events that range from life experiences on a life events inventory (Stromwall *et al.*, 2007) to a transgression that was committed by another person (e.g., Talwar, Lee, Bala, & Lindsay, 2004; Vrij, Akehurst, Soukara, & Bull, 2002) to being touched by a stranger (Edelstein, Luten, Ekman, & Goodman, 2006). Regardless of event, observers have difficulty determining the veracity of children's statements and their decisions appear to vary from study to study.

The target situations used in these studies have been criticized as being too mundane and lacking relevance to real legal cases (Vrij, Akehurst, Brown, & Mann, 2006). As criminal events are often highly salient and can be traumatic for children, research assessing perceptions of event veracity should attempt to observe children's lying behaviors in similarly stressful events. The target event needs to be salient and unique and should involve a high degree of emotion, as these factors could predict children's memory when providing details about the event (Bauer, 2006; Peterson, 2012). In support of this perspective, researchers have found impressively detailed long-term memories in children when they are interviewed about stressful events that they have previously experienced, although clear age differences are found, with the reports of older children longer and more coherent. The types of events assessed have included injuries (Peterson & Whalen, 2001), devastating natural disasters (Fivush *et al.*, 2004), and a painful medical procedure involving genital contact [the voiding cystourethrogram fluoroscopy (VCUG)] (Ornstein *et al.*, 1995; Quas *et al.*, 1999).

Supporting criticism of the events often used in judging children's veracity, using linguistics software to assess differences in the content of stressful versus non-stressful true and false stories, Brunet *et al.* (2013) found that true and false non-stressful stories could be better differentiated than true and false stories of victimization. Similarly, Saykaly *et al.* (2013) noted that lay judges could better recognize children's true non-stressful stories but were less accurate in assessing the veracity of stressful true stories and both their true and false stressful stories. Hence, it is possible that lay judges may differ in their ability to judge event veracity if they are assessing recall of a highly stressful experience, because children provide more details and provide stories that are contextually different (Brunet *et al.*, 2013; Saykaly *et al.*, 2013) about such events.

Factors that Potentially Influence Veracity Decisions

In discussing their findings about the believability of children who had or had not peeked, Crossman and Lewis (2006) suggested that the tendency to label children as lying may be an artifact of the short statements provided by children. Had children told longer stories, adults might have been more likely to believe them, a finding that has been documented in other studies (Orcutt *et al.*, 2001; Talwar & Lee, 2002). Differences in the amount of information children give in the lies they tell are substantial. Some studies (e.g., Crossman & Lewis, 2006) simply have children answer a small number of yes/no questions that adults rate for veracity while other studies have children provide full narratives. Some argue that because narratives have more detail than do answers to direct questions, narratives allow adults to more accurately judge veracity (Buck, Warren, Betman, & Brigham, 2002; Blandon-Gitlin, Pezdek, Rogers, & Brodie, 2005). Contrasting this, others argue that even with narratives, the few details typically provided by children make it difficult for adults to identify signs of deception (Lyon, Malloy, Quas, & Talwar, 2008). Nevertheless, differences in detail are a possible explanation for discrepancies in decisions made by judges in past studies.

The ability to assess veracity may also be influenced by the amount of preparation put into the lie (Stromwall *et al.*, 2007). Stromwall *et al.* (2007) found that when children were given a chance to prepare their own lies, adults only judged 47.6% of statements correctly. This was in contrast to the adults' ability to judge 55% of statements correctly when children were not given a chance to prepare their lies. In this study, children prepared their own lies. It is highly plausible that adults' coaching of children would have a similar negative effect, but this possibility has rarely been assessed.

Discussing the circumstances behind a witnessed or experienced crime may not be easy for a child. For example, children who were sexually abused may need to be prompted by adults to disclose information that they do not feel comfortable revealing (Hershkowitz, Lanes, & Lamb, 2007). Children in this position often feel ashamed to disclose details about the event, as they are anxious about the resulting responses. Adult coaching could be helpful for these children. In fact, Pezdek *et al.* (2004) have shown that in recalling a VCUG, children who had previously discussed the event with a parent had stories that were more believable than children who had not. It would be problematic, though, if adults were coaching children about information that did not happen. It is possible that adults going through custody battles, for example, might coach their child to lie in an effort to obtain custody of the child.

Tate, Warren, and Hess (1992) claim that coaching to lie does not work well with children. When the researchers coached children aged 2.6–8 years of age to lie about playing with a target toy, only seven out of 20 children were able to lie throughout the conversation they had with another person. Conversely, Lyon *et al.* (2008), who coached children to lie about a non-occurring event – playing with a toy house – found an opposite result. In that study, children aged 4–7 years who were coached to lie about playing with the toy house were highly successful at recounting details that they had been told about during the coaching session. One conclusion made in both studies was that older children were better at sustaining coached lies than younger children. Nevertheless, adults' ability to make veracity decisions was not incorporated into Tate *et al.*'s (1992) and Lyon *et al.*'s (2008) studies. Only recently have the effects of coaching been investigated in research assessing children's veracity.

In an attempt to assess the influence of coaching, Talwar, Lee, Bala, and Lindsay (2006) had 4- to 7-year-old children tell either a true or a false story about an event (e.g., a fishing trip). Prior to telling the story, children discussed it with their parents on multiple occasions. Talwar *et al.* (2006) found that the majority of children were judged as telling the truth (72.5%) regardless of the actual veracity of their stories. In fact, consistent with past research, the actual accuracy rate of judgements was 51%. The judges were no better at judging veracity as a whole, but their mistakes were due to an increase in the believability of children who were coached to lie.

Vrij, Akehurst, Soukara, and Bull (2002, 2004) found similar effects when they assessed veracity judgements of children who were prepared regarding how to tell a lie. In these studies, trained assessors were asked to use criterion-based content analysis (CBCA), a procedure used extensively in some European jurisdictions that assesses narratives for the presence of criteria commonly found in true stories, to determine whether children were telling the truth or a lie about either the children themselves or a researcher erasing notes from a chalkboard. Vrij *et al.* (2002) suggested that if people are determined to be convincing, they (or their lawyer) might research techniques that could be used to discriminate truth-telling from lie-telling. Prior to children's reports, Vrij *et al.* (2002, 2004) taught half of the children about details that would be looked for in a CBCA assessment. Children who were lying and not prepared with the CBCA information received lower CBCA scores, i.e., they were judged as less truthful, than children who were telling the truth. In contrast, children who were lying but were prepared with the CBCA information could not be distinguished from children who were telling the truth. However, the youngest children (5- and 6-year-olds) in Vrij *et al.*'s (2002) study could not be successfully coached to use the CBCA criteria.

Overall, the literature suggests that coaching presents a new challenge for adults. Discussions with adults appear to make children more credible. The specific content of the coaching can also influence the lies children are able to tell (Talwar, Murphy, & Lee, 2007b). For example, Talwar *et al.* (2007b) had parents coach 3-year-olds to lie about liking an undesirable gift (a bar of soap) and found that parents who gave more detailed instructions about lying had children who told more detailed lies. Whether or not children can be coached to lie about mundane events such as liking a bar of soap for a gift is a far cry from the more extensive and often highly salient events that are typical in forensic situations. In the present study, the influence of coaching on

both true and false accounts of a more complex and more stressful event (experiencing an injury severe enough to require emergency room treatment) is examined to determine whether similar effects are observed.

Information Medium

Traditionally, information presented in a courtroom is given verbally, with the witness present in the room with the jurors. This allows jurors to consider both the verbal information presented and the non-verbal behavior of the witness. However, this may not be the ideal way to assess the veracity of children's statements. In a direct comparison of the non-verbal behavior shown by 5- to 6-year-old children and adults, Vrij *et al.* (2004) found that children were much more active, moving nearly twice as much as the adults, which is often interpreted by observers as nervousness about lying. Vrij and colleagues believe that the best way to evaluate child truthfulness is to consider solely the content of children's stories. Others have found that artifacts of videotaping, such as camera angles, influence judgments of lying (Landstrom & Granhag, 2008).

Studies assessing how adults make decisions about veracity suggest that adults judge the verbal content more when judging children's veracity than that of adults (Landstrom, Granhag, & Hartwig, 2007; Stromwall & Granhag, 2007). These findings suggest that an analysis of the transcripts of children's statements may be more effective when assessing children's lie-telling than an analysis of audiovisual statements. In keeping with these suggestions, the present study used transcripts of children's interviews. Although some may see this as a limitation, transcripts of pre-trial interviews are often used by lawyers, prosecutors, and expert witnesses, and transcripts of trial testimony are sometimes used by jurors during deliberation.

The Present Study

In the previously mentioned coaching studies (Talwar *et al.*, 2007b; Vrij *et al.*, 2002, 2004), children were coached about relatively mundane events. In contrast, children in the present study were coached to discuss a stressful event. In addition, because age is often an important determinant of children's interview performance (Peterson, 2012), which in turn affects adults' expectations, we wanted to determine whether the influence of coaching was the same across children of varying ages. Children aged 5–7 years provide sparser narratives of events and have more difficulty producing coherent accounts than do older children (Peterson, 2012). Thus, coaching may not be as beneficial for them, particularly when coached to lie, as they need to maintain consistency and this may be difficult across a relatively lengthy interview. Older children, on the other hand, typically are better at maintaining consistency. Thus, coaching may substantially affect observers' perceptions of their veracity.

In the current study 5- to 14-year-old children lied or told the truth about an injury that was severe enough to warrant hospital emergency room treatment. Half of the children were coached by a parent to help them prepare their stories. Transcripts of these stories were presented to lay judges who were asked to assess their veracity and to indicate their confidence in that decision.

The following hypotheses were developed:

1. In keeping with past literature, lay judges would perform at levels similar to chance when judging the veracity of children's statements.

2. Lay judges would have more difficulty judging coached statements than judging uncoached statements. Past research has shown that coached statements are more believable than uncoached statements (Talwar et al., 2007b; Vrij *et al.*, 2002, 2004).
3. Lay judges would be more accurate in judging younger children's coached statements than those of older children.
4. Lay judges would have the most difficulty judging coached lies but would be more accurate in judging coached truths.
5. There would be no relationship between confidence and accuracy.
6. Judges would be more confident in their decisions regarding coached statements.

METHOD

Participants

A total of 1,074 undergraduate students (mean age, $M = 22$ years, range 17–72) were recruited as lay judges in the study: 329 males, 733 females and 12 participants who did not report gender. A draw for \$50 was used as an incentive to participate in the study. Most participants were Caucasian and of western European descent.

Materials

The materials were a collection of interviews with 96 children who described a target event about experiencing an injury. All interviews were audio recorded before being transcribed verbatim, with the exception of deleting any identifying information. The children interviewed were divided into three age groups: The “younger group” consisted of 5- to 7-year-olds ($n = 32$, $M = 5.9$, $SD = 0.98$, 16 females), the “middle group” consisted of 8- to 10-year-olds ($n = 32$, $M = 8.7$, $SD = 0.80$, 16 females), and the “older group” consisted of 11- to 14-year-olds ($n = 32$, $M = 12.5$, $SD = 1.3$, 16 females). For each age group, children were divided into four preparation conditions: true-coached, true-uncoached, false-coached and false-uncoached. Equal numbers of children were assigned to all preparation conditions within each age group. Thus, the study had a 3 (age) \times 2 (report: true vs. fabricated) \times 2 (coached vs. uncoached) design, with eight children (four females) in each cell.

The true-uncoached reports were randomly chosen from interviews previously conducted with children who had sustained real injuries and had been recruited at the emergency room of a children's hospital, the only children's hospital in the province and the only hospital treating children within a 100 mile radius (see Peterson, 2010). Additional children were recruited for the true-coached reports; these children were coached by their parents prior to being interviewed. All true accounts were corroborated by a parent. The fabricated reports were prepared by children who had never had a comparable injury requiring emergency room treatment (this was verified by parents). Some of the children in this condition were recruited through a childcare center and some were recruited through friends' and students' connections. These children were asked to fabricate a story about an injury that was similar to those sustained by the injured children. Parents were asked to practice their child's story with him/her at least once a day for the 4 days leading up to the interview, for approximately 10–20 minutes per session. The instructions for parents of children who were coached (both

those actually injured and those fabricating reports) are presented in the Appendix. Parents informally acknowledged using the coaching instructions but no data were collected regarding the actual coaching sessions.

During the interviews, all children were initially asked to provide free recall details about their injury as well as details about the subsequent hospital treatment. They were then asked a series of open-ended questions (e.g., “Where were you when it happened?”; “Who was there?”; “What did they do?”), followed by direct questions when it was necessary to clarify some of the children’s statements. The same interview has been used in prior research (e.g., Peterson, 1999; Peterson & Bell, 1996; Peterson & Whalen, 2001). Children in the fabricated condition had the same interview as children in the truth condition, except that those in the fabricated condition were told at the beginning of the interview to “try to fool other people” into believing that the injury had really happened to them.

In order to determine the similarities across interviews, the lengths of the children’s accounts (using word counts) were compared using 3 (age) \times 2 (true versus false) \times 2 (coached or not coached) between-subjects ANOVAs. There was an effect of age [$F(2, 83) = 6.46, p = 0.002, \eta_p^2 = 0.14$]. There was no difference in the word counts provided by 5- to 7-year-olds ($M = 256.93, SD = 141.41$) and 8- to 10-year-olds ($M = 292.48, SD = 155.88$; mean difference = 35.55, $p = 0.40$, 95% CI [-119.12, 48.03]). However, children in both the age ranges 5–7 and 8–10 years provided significantly fewer words than 11- to 14-year-olds ($M = 400.98, SD = 216.80$; mean difference = 144.05, $p = 0.001$, 95% CI [61.04, 227.05], and mean difference = 108.50, $p = 0.012$, 95% CI [24.92, 192.07] respectively). There was also an effect of coaching [$F(1, 83) = 0.94, p = 0.002, \eta_p^2 = 0.11$]. Coached children ($M = 370.75, SD = 198.32$) provided significantly higher word counts than uncoached children ($M = 262.85, SD = 150.81$, 95% CI [39.82, 175.98]). There was no effect of event veracity or any two- or three-way interactions between children’s age, event veracity, and whether or not children were coached on children’s provision of information.

Procedure

Each lay judge was randomly assigned to receive one of the 96 interview transcripts and was asked to complete a short questionnaire. The questionnaire asked participants to indicate whether they thought the children were lying or telling the truth, and to rate their level of confidence in their judgments on a five-point scale from 1 (= “not at all confident”) to 5 (= “very confident”). In addition, information was collected on their age and gender. Participants could take their package home but were asked to not talk to others about the transcript they had been given. They were to return their package in their next class or drop it off at a box in the Psychology Department office.

RESULTS

Overall Veracity Assessment

Table 1 provides a breakdown of the proportion of accurate decisions that were made by lay judges in each condition. There was no difference in the ability to judge the accuracy of boys versus girls. Thus gender was not considered further. One-sample *t*-tests

Table 1. Accuracy rates (%) for participants in the various conditions

Condition	Age group			
	5–7 years	8–10 years	11–14 years	All ages
Uncoached-false	60.19*	53.92	51.35	55.14
Uncoached-true	45.83	50.50	48.00	48.35
Coached-false	65.06**	35.16**	32.73**	42.96*
Coached-true	53.52	71.15**	49.32	56.63
All conditions	56.89*	50.58	44.92*	50.47

Note: **indicates significantly different from chance (50%) at $p < 0.01$; *indicates significance at $p < 0.05$ in one-sample t -tests comparing proportion accurate to chance level accuracy.

were completed to determine whether the proportion of accurate decisions in a particular condition were significantly different from chance (50%) accuracy. The lay judges' overall accuracy rate for detecting truths and lies was 50.5%, which is statistically no different than chance [$t(1073) = 0.31, p = 0.76$]. When liars were considered specifically, the lay judges' accuracy rate was 49.6%, which is again no different than chance [$t(598) = -0.20, p = 0.84$]. Similarly, when truth-tellers were considered, the accuracy rate was 51.6%, which is no different than chance [$t(474) = 0.69, p = 0.49$]. These findings are consistent with Hypothesis 1.

Effects of Children's Age, Event Veracity, and Coaching

To determine whether the accuracy of veracity decisions could be predicted from child age group, event veracity, and whether or not the children were coached, as well as the interactions amongst these predictors, a series of binary logistic regression analyses were completed with accuracy as the outcome variable. In order to control the order of entry in which interactions were to be considered, the enter method was used. First, the main effects of age, veracity, and coaching were examined. The overall model was significant [model $\chi^2(3) = 12.19, p = 0.007, R_N^2 = 0.02$]. Prediction accuracy was improved from 50.5% ($-2LL = 1,488.79$) using just the constant to 56.7% with the three predictor variables added ($-2LL = 1,476.59$). The analyses revealed that children's age was a significant predictor of accuracy (Wald = 10.31, $p = 0.001$). As shown in step 1 in Table 2, the odds ratio (OR) associated with children's age suggests the younger the age group a child is in, the more likely participants were to be accurate in their veracity decision. In contrast to Hypothesis 2, there was no effect of coaching on the overall accuracy of veracity judgements.

Next, potential two-way interactions were assessed. In order to do this, the main effects were entered on the first step of the regression analysis and then the interaction between two variables was included in a second step. When the interaction between child age and event veracity was included in a regression model after child age, event veracity, and whether or not the child was coached were already in the model, the incremental change was significant [$\chi^2(1) = 5.76, p = 0.02 (-2LL = 1470.83)$]. Follow-up regressions were then completed to assess this interaction. Event veracity influenced accuracy for the adults who read transcripts from 5- to 7-year-olds (Wald = 5.62, $p = 0.02$). More specifically, adults' decisions were more likely to be accurate if they read

Table 2. The logistic regression model predicting decision accuracy

	<i>B</i> (SE)	95% CI for odds ratio		
		Lower	Odds ratio	Upper
Step 1				
Constant	0.47 (0.17)			
Child age group	-0.24** (0.08)	0.68	0.79	0.91
Event veracity	0.09 (0.12)	0.86	1.10	1.40
Coaching	-0.07 (0.06)	0.83	0.94	1.06
Step 2				
Constant	0.80 (0.22)			
Child age group	-0.41** (0.10)	0.55	0.67	0.81
Event veracity	-0.63 (0.34)	0.27	0.53	1.03
Coaching	-0.09 (0.18)	0.78	1.10	1.55
Child age group × event veracity	0.37* (0.15)	1.17	1.50	1.91
Child age group × coaching	-0.16* (0.08)	0.74	0.85	0.99
Coaching × event veracity	0.40** (0.13)	1.17	1.50	1.91

Note: **, $p < 0.01$; *, $p < 0.05$. CI, confidence interval.

false accounts than if they read true accounts from 5- to 7-year-olds (OR = 0.59, 95% CI [0.38, 0.91]). Similarly, event veracity influenced accuracy for the adults who read transcripts from 8- to 10-year-olds (Wald = 4.92, $p = 0.03$). In this case, adults' decisions were more likely to be accurate if they read true accounts than if they read false accounts (OR = 1.63, 95% CI [1.06, 2.51]). Event veracity did not influence the accuracy of decisions made by adults who read transcripts from 11- to 14-year-olds (Wald = 1.29, $p = 0.26$).

Partially supporting Hypothesis 3, when the interaction between child age and whether or not the child had been coached was included in a regression model after age, event veracity, and whether or not the child had been coached were already in the model, the incremental change was significant [$\chi^2(1) = 4.32$, $p = 0.04$ (-2LL = 1,472.27)]. Once again, follow-up regressions were completed to assess the interaction. When a child was not coached the age of the child did not influence the accuracy of adults' veracity decisions, Wald = 0.70, $p = 0.40$. However, when the transcript being read was from a coached child, age influenced the accuracy of veracity decisions, Wald = 12.88, $p < 0.001$. The accuracy of veracity decisions was higher if the child was younger than if the child was older (OR = 0.67, 95% CI [0.54, 0.83]).

When the interaction between whether or not the child had been coached and event veracity was included in a regression model after age, whether or not the child had been coached and event veracity were already in the model, the incremental change was significant [$\chi^2(1) = 9.93$, $p = 0.002$ (-2LL = 1,466.67)]. Supporting Hypothesis 4, follow-up regression analyses showed when the event was false, coaching affected accuracy (Wald = 7.93, $p = 0.005$). Judges made more errors in judging a coached child than in judging a child who was not coached (OR = 0.79, 95% CI [0.67, 0.93]). In contrast, when the event was true, coaching did not influence judgement accuracy (Wald = 3.08, $p = 0.08$).

The possible existence of a three-way interaction between child age group, event veracity and whether or not the children were coached was assessed by entering the three associated main effects on the first step of a regression analysis followed by the three associated two-way interactions on the second step, and the three-way interaction on

the third step of the analysis. The incremental change when adding the three-way interaction after the main effects and two-way interactions were already in the model was not significant [$\chi^2(1) = 1.31, p = 0.25$ ($-2LL = 1,455.00$)], showing that the three-way interaction between these predictors did not influence the accuracy of veracity decisions.

After determining the presence of three significant interactions, a final regression analysis was completed with the main effects on the first step of the model and the three two-way interactions on the second step of the model. The overall model was significant [model $\chi^2(6) = 32.48, p < 0.001, R_N^2 = 0.04$]. Prediction accuracy was improved from 50.5% using just the constant ($-2LL = 1,488.78$) to 56.7% with the three predictor variables and the three two-way interactions between the predictors added ($-2LL = 1,456.30$). As seen in the second step of the model shown in Table 2, the main effect of children's age group held (Wald = 15.98, $p < 0.001$). The OR again showed that the younger the child, the more accurate participants were in their decisions. Finally, the three two-way interactions continued to significantly influence the accuracy of veracity decisions in the overall model.

Confidence

A Pearson correlation was performed to examine the relationship between participants' confidence judgments on a scale of 1 to 5 ($M = 3.62, SD = 0.711$) and their decision accuracy (correct vs. incorrect). There was no relationship between confidence and accuracy ($r = -0.04, n = 1073, p = 0.18$).

Next a 3 (age group) \times 2 (event veracity) \times 2 (coached or not) ANOVA was completed to determine whether there were any differences in confidence across the various conditions. There was no difference in participants' confidence levels when judging children in each age group or when judging true versus false statements, but participants were more confident in their responses when judging coached compared with uncoached statements [$F(1, 1061) = 4.82, p = 0.03, \text{partial } \eta^2 = 0.01$].

DISCUSSION

In the present study, lay judges' ability to judge the veracity of children's accounts of an injury severe enough to require emergency room treatment was examined. Not surprisingly, given the accuracy rates found in other similar studies (e.g., Stromwall *et al.*, 2007; Talwar *et al.*, 2006), adults had difficulty making an accurate veracity decision, regardless of whether a child was telling the truth or a lie (Hypothesis 1). However, the accuracy of the decision made could be partly predicted from a combination of three factors – children's age, event veracity and whether or not children had been coached by a parent.

A main effect of age shown in the initial logistic regression analysis indicated that age better predicts accuracy for younger children. As seen in Table 1, younger children (aged 5–7 years) were relatively incapable of successfully maintaining a lie (Hypothesis 3). This may be partly explained by the fact that younger children provided shorter accounts (i.e., fewer words) than older children but given that both 5- to 7- and 8- to 10-year-olds provided accounts of similar length, if the effect of

age was simply a function of the number of words provided, we should see this with both age groups. Furthermore, as demonstrated in Table 1 and in later analyses, the predictability afforded by knowing age varied according to event veracity and whether or not the children had been coached. Adults better assessed 5- to 7-year-olds' false accounts than their true accounts. In contrast, adults better assessed 8- to 10-year-olds' true accounts than their false accounts. This again suggests that the differences in accuracy are not simply a function of the length of the children's accounts. There was no difference in the accuracy of judgements for 11- to 14-year-olds regardless of whether the event was true or false.

Coached children, regardless of age, provided longer accounts than uncoached children. Despite this, coached 5- to 7-year-olds were still not believed. Coaching significantly increased the likelihood of adults' believing the accounts of 8- to 10-year-olds regardless of whether the child was truthful or lying. Similarly, 11- to 14-year-olds who were coached were more likely to be seen as telling the truth than those who were not coached. In particular, 11- to 14-year-olds who were coached were believed at a higher rate than chance levels when they were telling a lie.

Past research has shown conflicting results with respect to how successfully adults can coach children. In their assessments of children who were coached to talk about play activities, Tate *et al.* (1992) and Lyon *et al.* (2008) found it difficult to successfully coach young children to describe activities using details provided by an adult. Similarly, Vrij *et al.* (2002) noted their difficulty in successfully coaching young children to use CBCA criteria. Contrasting this, parents in Talwar *et al.*'s (2006) study were successful in coaching 4- to 7-year-olds to tell believable lies. The notable difference among these studies was the person who instigated the coaching – the Talwar *et al.* (2006) study used parents rather than strangers. This appeared to suggest that parents, but not necessarily strangers, could coach a young child to tell a story. However, in contrast to the Talwar *et al.* (2006) study and similar to the findings of Tate *et al.* (1992) and Lyon *et al.* (2008), parents in the present study were not successful in coaching their 5- to 7-year-olds to lie. In fact, in the present study, lay judges performed at above-chance levels when evaluating both the coached and uncoached lies of 5- to 7-year-olds.

Vrij (2000) has suggested through his cognitive load hypothesis that deception is more difficult to maintain than truthfulness, because lying requires more effort to maintain a consistent report. Younger children have more limited cognitive resources available, possibly explaining why their lies were less convincing in this study. Rather than helping, coaching may have hindered these children, as they now had more things to consider when providing their story. The children were left with the task of balancing their story, suggestions parents had provided, worry as to the believability of the story, and worry about appropriately responding to the questions they were being asked. Providing a believable story proved difficult and 5- to 7-year-olds who were lying provided contradictory or unbelievable information regardless of whether they were coached or not. Children in Talwar *et al.*'s (2006) study faced a similar task, but they discussed mostly positive events. Thus, there seem to be differences in the success of parental coaching as a function of the event children are being coached about.

Another notable difference is that the lay judges in Talwar *et al.*'s study were shown a video of children being interviewed and then cross-examined about the event in question. In contrast, in the present study, lay judges were presented with written transcripts of an interview. Differences between studies on the success of coaching may

depend on the medium of information, but this would be impossible to assess without a direct comparison of video and transcript presentation formats.

As noted in Table 1, the 8- to 10-year-olds who were coached were believed regardless of event veracity. Thus, children in this group were more successful at maintaining consistency in their coached lies when fabricating, unlike their younger counterparts. They also were successful at incorporating coached suggestions and information in order to make their true accounts more believable. Vrij *et al.* (2004) did include some 8-year-olds in their study assessing whether children could incorporate CBCA criteria into event descriptions, but did not find a similar effect of coaching. In their study, the 8-year-olds were paired with 6- and 7-year-olds in the overall analyses, supporting the idea that when younger children are evaluated, they have problems maintaining a lie. Added to this, these children were asked to use the CBCA criteria. It could be argued that being asked to use criteria that they have never had previous experience with would add to the cognitive load these children are faced with, making the task of developing and maintaining a convincing lie more difficult even for these slightly older children. Furthermore, a stranger rather than a parent was used as the person doing the coaching, which may be an important distinction for children this age. The present study is the first known study to demonstrate how highly sensitive children in this age group are to coaching. Future research should assess the circumstances under which such powerful effects of coaching occur.

The findings were again different for the 11- to 14-year-olds who, when coached to lie, were believed at greater than chance levels (Hypothesis 3). This fits with the findings of Vrij *et al.* (2002, 2004), who noted that older children who were coached to use CBCA criteria provided stories that were similar in believability to true stories. Interestingly, correct identification of true-coached reports from children of this age did not differ from chance levels, as shown in Table 1. Perhaps when older children were coached in telling about an injury that they had previously experienced, their story turned out to be too perfect, thus making lay judges doubt the veracity of the story. An alternative explanation may be that parents were not as engaged in coaching children to tell a factual story as they would have been if their child needed to fabricate one. A limitation of the present study is that the actual content of coaching was not assessed. The older children who were asked to recount true stories could probably be assumed to sustain a well-versed memory about their injury, and thus thorough coaching was not thought to be needed by parents to help these children. In other words, perhaps parents were less motivated to help these children develop coherent truthful accounts.

Crossman and Lewis (2006) suggested that the children in their study were rated as lying more often than telling the truth, possibly because of the short statements they used. Similarly, Lyon *et al.* (2008) indicated that the few details generally provided by children when describing something make it difficult to identify signs of deception. In the Talwar *et al.* (2006) study where adults were given ample opportunity to assess children's deception, there was a truth bias. As in the Talwar *et al.* (2006) study, in the present study, adults were provided with a full narrative from the children. This suggests that both the difficulties children of varying ages have in providing details and the nature of the events described may explain differences in judgements of veracity.

The present study is not without limitations. In the current study most of the children who fabricated stories about an injury presumably did not experience negative

emotions such as guilt or fear of punishment while telling these stories. Our instruction of asking the children “to try and fool” the interviewer was assumed to lessen the negative feelings that children might have initially felt. Also, parents’ coaching could have acted as an indicator that it is acceptable to fabricate a lie. Hence, children whose negative emotions were absent during the interview would be more eloquent at telling mendacious stories than children who sustained those negative feelings. Clearly, an actual courtroom setting does not operate in the same way as our interview session did. An additional limitation was that parents were given extensive hints about the sort of information that would be queried in their child’s interview. This could be seen as leading to more intensive coaching than would generally be expected; however, as suggested by Vrij *et al.* (2002), if individuals are motivated enough, they may go to far greater measures (e.g., searching for CBCA criteria on the internet) to coach a child. Talwar *et al.* (2007b) suggest that children differ in their ability to maintain coached statements as a consequence of more versus less detailed coaching. In future studies, it might help to tape the coaching completed by parents so that this might be assessed. Finally, some of the lay judges in our study took the transcripts home. It is possible then that even though they were instructed not to discuss the transcript with others, lay judges may have had outside help in coming to a decision.

Real-life events that lead to children’s engagement in the forensic arena differ from the event studied here in a host of important ways. The finding that lay judges in the present study were poor at discriminating false from true accounts is disturbing. This was found even though lay judges had at their disposal lengthy interviews about a complex event, rather than short statements. In addition, the event was highly salient and stressful, and children have been shown to have excellent long-term memory for details about these sorts of events (Peterson, 1999, 2012). The potential consequences of an inability to discern truthful reports are that if children who are telling the truth are not believed, children who are abused may continue to suffer; moreover, the system fails them and the fact that they are disbelieved adds considerably to their distress.

Just as important in the present study, parental coaching was effective with children who were at least 8 years of age, which has disconcerting implications for the judicial system, because it is not uncommon for parents (or lawyers) to coach child witnesses prior to testifying in court. The fact that coached lies can seem more convincing and believable than even true accounts that were provided by children without prior coaching suggests that determining the truthfulness of child witnesses may be even more difficult than previously thought. Considerable research has shown that adults are poor at discriminating between true and false statements. It is even more disconcerting if the reports that are most convincing are coached. In fact, in the present study, even though there was no relationship between confidence and accuracy (Hypothesis 5), lay judges were most confident in their assessments of coached statements (Hypothesis 6). These statements are convincing and those who hear them will be confident that they have made an accurate decision about the veracity of the statements. Thus, coaching may potentially have a strong and under-appreciated effect on the outcome of forensic investigations, pre-trial hearings, veracity assessments, and trials. Past research has demonstrated that police officers are more confident in their ability to make accurate veracity decisions than lay judges (DePaulo & Pfeifer, 1986). If police officers show a similar problem in identifying the veracity of coached accounts, they may choose to actively pursue claims where an investigation is not needed or to ignore claims where an investigation is necessary. This could lead to turmoil in the lives of

those involved, regardless of how an investigation turns out. How much one believes a child witness should not be dependent upon how well coached that child is. Clearly, considerably more research needs to be conducted to explore the impact of coaching on children's believability.

REFERENCES

- Bauer, P. J. (2006). Event memory. In R. M. Lerner (Ed.), *Handbook of child psychology: Vol 2, cognition, perception, and language (6th ed.)*. (pp. 373–425). Hoboken, NJ US: John Wiley & Sons Inc.
- Blandon-Gitlin, I., Pezdek, K., Rogers, M., & Brodie, L. (2005). Detecting deception in children: An experimental study of the effect of event familiarity on CBCA ratings. *Law and Human Behavior, 29*, 187–197. doi:10.1007/s10979-005-2417-8.
- Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*, 214–234. doi:10.1207/s15327957pspr1003_2.
- Brunet, M. K., Evans, A. D., Talwar, V., Bala, N., Lindsay, R. C. L., & Lee, K. (2013). How children report true and fabricated stressful and non-stressful events. *Psychiatry, Psychology and Law, 20*, 867–881. doi:10.1080/13218719.2012.750896.
- Buck, J. A., Warren, A. R., Betman, S. I., & Brigham, J. C. (2002). Age differences in criteria-based content analysis scores in typical child sexual abuse interviews. *Applied Developmental Psychology, 23*, 267–283. doi:10.1016/S0193-3973(02)00107-7.
- 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.
- DePaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. J., & Muhlenbruck, L. (1997). The accuracy-confidence correlation in the detection of deception. *Personality and Social Psychology Review, 1*, 346–357. doi:10.1207/s15327957pspr0104_5.
- DePaulo, B. M., & Pfeifer, R. L. (1986). On-the-job experience and skill at detecting deception. *Journal of Applied Social Psychology, 16*, 249–267. doi:10.1111/j.1559-1816.1986.tb01138.x.
- 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.
- Fivush, R., Sales, J., Goldberg, A., Bahrlick, L., & Parker, J. (2004). Weathering the storm: Children's long-term recall of Hurricane Andrew. *Memory, 12*, 104–118. doi:10.1080/09658210244000397.
- Hershkowitz, I., Lanes, O., & Lamb, M. E. (2007). Exploring the disclosure of child sexual abuse with alleged victims and their parents. *Child Abuse and Neglect, 31*, 111–123. doi:10.1016/j.chiabu.2006.09.004.
- Landstrom, S., & Granhag, P. A. (2008). Children's truthful and deceptive testimonies: How camera perspective affects' adult observers' perception and assessment. *Psychology, Crime & Law, 14*, 381–396. doi:10.1080/10683160701580107.
- Landstrom, S., Granhag, P. A., & Hartwig, M. (2007). Children's live and videotaped testimonies: How presentation mode affects observers' perception, assessment and memory. *Legal and Criminological Psychology, 12*, 333–347. doi:10.1348/135532506X133607.
- Leach, A. M., Lindsay, R. C. L., Koehler, R., Beaudry, J. L., Bala, N. C., Lee, K., ... Talwar, V. (2009). The reliability of lie detection performance. *Law and Human Behavior, 33*, 96–109. doi:10.1007/s10979-008-9137-9
- 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 truth/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. A. (2008). Coaching, truth indication, and young maltreated children's false allegations and false denials. *Child Development, 79*, 914–929. doi:10.1111/j.1467-8624.2008.01167.x.
- Orcutt, H. K., Goodman, G. S., Tobey, A. E., Batterman-Faunce, J. M., & Thomas, S. (2001). Detecting deception in children's testimony: Factfinders' abilities to reach the truth in open court and closed-circuit trials. *Law and Human Behavior, 25*, 339–372. doi:10.1023/A:1010603618330.
- Ornstein, P. A., Baker-Ward, L., Myers, J., Principe, G. F., & Gordon, B. N. (1995). Young children's long-term retention of medical experiences: Implications for testimony. In W. Schneider (Ed.), *Memory performance and competencies: Issues in growth and development*. (pp. 349–371). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.
- Peterson, C. (1999). Children's memory for medical emergencies: 2 years later. *Developmental Psychology, 35*, 1493–1506. doi:10.1037/0012-1649.35.6.1493.

- Peterson, C. (2010). "And I Was Very Very Crying": Child Self-Descriptions of Distress as Predictors of Recall. *Applied Cognitive Psychology*, 24, 909–924. doi:10.1002/acp.1636.
- Peterson, C. (2012). Children's autobiographical memories across the years: Forensic implications of childhood amnesia and eyewitness memory for stressful events. *Developmental Review*, 32, 287–306. doi:10.1016/j.dr.2012.06.002.
- Peterson, C., & Bell, M. (1996). Children's memory for traumatic injury. *Child Development*, 67, 3045–3070. doi:10.2307/1131766.
- Peterson, C., & Whalen, N. (2001). Five years later: Children's memory for medical emergencies. *Applied Cognitive Psychology*, 15, S7–S24. doi:10.1002/acp.832.
- Pezdek, K., Morrow, A., Blandon-Gitlin, I., Goodman, G. S., Quas, J. A., Saywitz, K. J., ... Brodie, L. (2004). Detecting deception in children: Event familiarity affects criterion-based content analysis ratings. *Journal of Applied Psychology*, 89, 119–126. doi:10.1037/0021-9010.89.1.119
- 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., Bidrose, S., Pipe, M.-E., Craw, S., & Ablin, D. S. (1999). Emotion and memory: Children's long-term remembering, forgetting, and suggestibility. *Journal of Experimental Child Psychology*, 72, 235–270. doi:10.1006/jecp.1999.2491.
- Saykaly, C., Talwar, V., Lindsay, R. C. L., Bala, N., Lee, K., ... & Nugent, S. M. (2013). Adults ability to detect deception of stressful and non-stressful stories of children. *Psychology Crime and Law*, 19, 865–879. doi:10.1080/1068316X.2012.700311
- Stromwall, L. A., & Granhag, P. A. (2007). Detecting deceit in pairs of children. *Journal of Applied Social Psychology*, 37, 1285–1304. doi:10.1111/j.1559-1816.2007.00213.x.
- Stromwall, L. A., Granhag, P. A., & Landstrom, S. (2007). Children's prepared and unprepared lies: Can adults see through their strategies? *Applied Cognitive Psychology*, 21, 457–471. doi:10.1002/acp.1288.
- Talwar, V., Lee, K., Bala, N., & Lindsay, R. C. L. (2004). Children's lie-telling to conceal a parent' transgression: Legal implications. *Law and Human Behavior*, 28, 411–435. doi:10.1023/B:LAHU.0000039333.51399.f6.
- Talwar, V., Gordon, H. M., & Lee, K. (2007a). Lying in the elementary school years: 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. (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., Bala, N., & Lindsay, R. C. L. (2006). Adults' judgments of children's coached reports. *Law and Human Behavior*, 30, 561–570. doi:10.1007/s10979-006-9038-8.
- Talwar, V., Murphy, S. M., & Lee, K. (2007b). White lie-telling in children for politeness purposes. *International Journal of Behavioral Development*, 31, 1–11. doi:10.1177/0165025406073530
- Tate, C. S., Warren, A. R., & Hess, T. M. (1992). Adults' liability for children's 'lie-ability': Can adults coach children to lie successfully? In M. Putnick (Ed.), *Cognitive and social factors in early deception*. (pp. 69–87). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.
- Tye, M. C., Amato, S. L., Honts, C. R., Devitt, M. K., & Peters, D. (1999). The willingness of children to lie and the assessment of credibility in an ecologically relevant laboratory setting. *Applied Developmental Science*, 3, 92–109. doi:10.1207/s1532480xads0302_4.
- Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice*. Chichester, UK: John Wiley & Sons.
- Vrij, A., Akehurst, L., Brown, L., & Mann, S. (2006). Detecting lies in young children, adolescents, and adults. *Applied Cognitive Psychology*, 20, 1225–1237. doi:10.1002/acp.1278.
- Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2002). Will the truth come out? The effect of deception, age, status, coaching, and social skills on CBCA scores. *Law and Human Behavior*, 26, 261–283. doi:10.1023/A:1015313120905.
- Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2004). Let me inform you how to tell a convincing story: CBCA and reality monitoring scores as a function of age, coaching, and deception. *Canadian Journal of Behavioural Science*, 36, 113–126. doi:10.1037/h0087222.

APPENDIX: INSTRUCTIONS FOR PARENTS OF CHILDREN WHO WERE COACHED

Before the interview, you and your child will practice his/her story together to try and make it more believable. This is meant to reflect actual situations where parents may

attempt to instruct their child to tell a specific lie to authorities such as the police. Consequently it is important that you directly help your child to work out the details of the story and the most convincing way to tell it. You and your child should practice at least once a day for the 4 days leading up to the interview, for approximately 10–20 minutes per session. The interviews tend to have some standard open-ended questions and here are some general points to consider elaborating in your story.

General situation:

- Where was it?
- What time of day was it?
- What season/weather was it?
- What activity was occurring when it happened?
- What were people dressed like?

Before incident:

- Who was around?
- What specifically was your child doing?
- What were others doing?
- What objects were nearby?
- Were there any arguments or other interactions occurring between people involved?

During incident:

- Who witnessed the incident?
- What did they see?
- What was the physical cause of injury?
- Was there any important dialogue?

Injury:

- Describe any pain, bleeding or discomfort.
- Visual/auditory description of the injury.

Immediately after incident:

- Reactions of your child or bystanders to injury including: dialogue, emotions, crying and attempts at treatment.

After incident:

- Who helped and how and why?
- How was help found?
- Description of immediate treatment of injury.

Aftermath:

- How were the parents found and involved?
- Was there any first aid?
- Describe travelling to the hospital, ie. who accompanied the child.

Also:

Don't worry if any details you include may seem unfamiliar to readers.

Thank you for your help in this study.