Psychiatry, Psychology and Law, 2018
VOL. 25, NO. 5, 789–805
https://doi.org/10.1080/13218719.2018.1478336

Children who are coached to lie: does linguistic analysis help in understanding why these children are so believable?

Kelly L. Warren[a], Carole Peterson[b] and Cassy C. Gillingham[b]

[a]*Psychology Program, Memorial University of Newfoundland, Corner Brook, Canada;* [b]*Department of Psychology, Memorial University of Newfoundland, St. John's, Canada*

In this study, the usefulness of linguistic analysis in determining the veracity of children's accounts is examined. The Linguistic Inquiry Word Count 2007 program was used to analyze 95 stories told by 5- to 14-year-olds who were telling the truth or a lie about the stressful experience of breaking a bone or requiring sutures for serious lacerations. Half of the children were coached by parents in preparing their story over the four days prior to giving their account. Differences emerged in the linguistic style used as a function of age, presence of coaching and event veracity. Very few linguistic categories emerged as significant predictors of event veracity, and the variables that did emerge were different depending upon the presence of coaching. Since in real-life situations one seldom knows a child's coaching history, these findings suggest that it is inappropriate to use linguistic analysis to assess the veracity of children's accounts.

Key words: child; coaching; deception; LIWC; parent; veracity.

Research assessing laypersons' ability to determine the veracity of children's stories has generally demonstrated an ability that is little different from chance (Stromwall, Granhag, & Landstrom, 2007; Talwar, Lee, Bala, & Lindsay, 2004; Vrij, Akehurst, Soukara, & Bull, 2002). Regardless of whether children are providing a lengthy or short account (Crossman & Lewis, 2006; Lyon, Malloy, Quas, & Talwar, 2008; Talwar & Lee, 2002), describing a stressful or non-stressful event (Talwar, Lee, Bala, & Lindsay, 2006; Warren, Dodd, Raynor, & Peterson, 2012; Warren et al., 2015) or telling prepared or unprepared truths or lies (Stromwall et al., 2007; Warren et al., 2012), people falter when asked to make an intuitive decision about the truthfulness of a story. Differences exist in the ability to make veracity decisions across these variables, but to date – despite the complexity or simplicity of the variables being examined – laypersons have rarely been shown to be able to judge children's event veracity with accuracy. The picture is even more problematic when people are asked to judge children's coached accounts (Talwar et al., 2006; Vrij, Akehurst, Soukara, & Bull, 2004; Vrij et al., 2002) – that is, when people are judging stories told by children who have been directed in preparing their true or false stories by a parent or some other adult.

The assessment of coached accounts is important as some studies indicate that false allegations of mistreatment and other crimes

Correspondence: Kelly Warren, Psychology Program, Grenfell Campus, Memorial University of Newfoundland, 20 University Drive, Corner Brook, NL, Canada A2H 5G4. Email: kwarren@grenfell.mun.ca

against children often originate with an adult (Bala, Lee, & McNamara, 2001; Trocme & Bala, 2005). Similarly, professionals who work with children in a criminal justice context sometimes indicate a belief that they have worked on cases in which a child was coached to provide a false account of an event (Faller, 2007). Recognizing the importance of examining coaching as potentially influencing the believability of children's accounts, a number of recent studies have explored differences in making veracity decisions when coaching is or is not involved (Talwar, Murphy, & Lee, 2007; Vrij et al., 2002, 2004; Warren et al., 2015). These studies indicate that coaching does not appear to have an impact when very young children are coached – and, in fact, that people often accurately detect lies told by young children (Vrij et al., 2002; Warren et al., 2015). However, the accounts given by older children who are coached to lie are believed at a higher rate than the accounts told by children who are telling unprepared lies, self-prepared lies or even, in some cases, the truth (Talwar et al., 2007; Vrij et al., 2002, 2004; Warren et al., 2012). The fact that coached lies may be more believable than true (uncoached) accounts is dismaying. The increased believability of coached accounts told by children occurs whether children are coached about a positive or a negative event, and whether coaching is being carried out by a parent or some other adult (Talwar et al., 2007; Vrij et al., 2002, 2004; Warren et al., 2015).

Given the difficulty that laypersons have in deciphering whether an account is true or false, in the adult literature some recent research has used linguistic software in an attempt to determine whether or not computer programs can be used to differentiate between truth and lies (e.g. Ali & Levine, 2008; Bond & Lee, 2005; Newman, Pennebaker, Berry, & Richards, 2003). The most commonly used program is Linguistic Inquiry and Word Count (LIWC), which is designed to analyze written and spoken accounts of an event on a word-by-word basis (Pennebaker, Francis, & Booth, 2001). Although LIWC was not designed to assess veracity, it has been used by some to make this distinction on the premise that telling a lie should require effort and consequently involve a different pattern of language usage (e.g. Newman et al., 2003). Researchers have had some, albeit limited, success in using LIWC to determine whether people are telling the truth or lying, and have identified some features that seem to differentiate true accounts from false ones (Hauch, Blandon-Gitlin, Masip, & Sporer, 2015).

More specifically, a recent meta-analysis (Hauch et al., 2015) which combined studies that assessed adults with studies that assessed children demonstrated that some but not all of the traditionally held assumptions about truth–lie distinctions are detected by computer-generated linguistic analyses. Consistent with expectations, at least some components analyzed using linguistic software suggest that liars experience a greater cognitive load than truth-tellers, as well as expressing more negative emotions and using fewer self-references and more other-references. Additionally, the meta-analysis indicates that truth-tellers use more cognitive processes than liars. In contrast to expectations, liars do not appear to be more uncertain than truth-tellers and in fact the opposite may be the case. The findings with respect to emotions (positive, negative or undefined) are somewhat contradictory when comparing truth-tellers to liars, and no differences were observed in relation to generalizing terms, the use of the passive voice and verb tenses. Notably, although studies that assessed children were combined with those that assessed adults in their meta-analysis, Hauch et al. (2015) indicate as a limitation of their study that the two groups should have been investigated separately. Thus, since the majority of studies included in their meta-analysis involved adults, it is unclear whether or not their findings pertain to children (Hauch et al., 2015).

In developing their own studies using linguistic analyses to evaluate the veracity of

stories, many individuals have used the model developed by Newman et al. (2003) in an attempt to determine whether or not the factors that predict veracity decisions are consistent across studies. In their original article, Newman et al. argue that it is important to develop a model of lie-telling which is independent of event content, variables that are infrequently used and variables that could be unique to a spoken or written account. They further argue that it is necessary to look at multiple linguistic cues rather than individual cues, as the latter prevents the formation of a multivariate profile of true versus false accounts. This left them with 29 LIWC categories, which they subsequently examined in the accounts of adults. Five variables were found to be potential indicators of event veracity for adults: those who were lying used fewer first-person singular pronouns, fewer third-person pronouns, more negative emotion words, fewer exclusive words and more motion verbs. These variables have since been assessed in several more recent studies that have assessed the veracity of both adults' and children's accounts (Bond & Lee, 2005; Brunet et al., 2013; Evans et al., 2012).

Despite a somewhat expansive literature using such software in research with adults, research using linguistic software to assess the veracity of children's accounts is limited, and – as suggested by Sim and Lamb (2013) – the usefulness of computerized language programs for examining children's event accounts is unclear and may possibly be inappropriate. In their study, Sim and Lamb assess the linguistic patterns exhibited in interviews with alleged child sexual abuse victims. Although ground truth could not be fully determined in these interviews, Sim and Lamb reason that it is important to determine whether or not an LIWC analysis indicates differences in the children's accounts as a function of other variables. In particular, the authors look at differences in language patterns according to the age and gender of the child, the child's relationship with the suspect, the type of abuse, the number of incidents and the type of interview. Group differences emerged across all of these variables for possible predictors previously identified by Newman et al. (2003) in the adult literature and for the LIWC categories that Sim and Lamb identify as being consistent with reality monitoring (RM) and criterion-based content analysis (CBCA) approaches to detecting deception (affective processes, perceptual processes, feel words, space words, time words and insight words). Sim and Lamb argue that additional research with children is crucial, as it appears that differences may exist between the language used by children versus adults when attempting to deceive others or to tell the truth.

A limited amount of research has been undertaken to assess linguistic differences in children's truthful versus deceitful accounts. For example, Evans et al. (2012) had parents coach their children (4- to 10-year-olds) to provide interviewers with a report of two games: one the children had played and one they had not played. A direct comparison of the information provided by the children in describing the games they had versus had not played indicated that the 6- to 7-year-olds provided significantly more sensory/perceptual terms and causation terms in their false accounts than in their true accounts. An examination of whether or not there were differences in the specific patterns of language usage across the true and false accounts showed that the combination of age group, use of sensory/perceptual processes, use of causation terms, use of words with six or more letters, use of second-person pronouns and use of tentative terms could correctly predict true or false accounts 66.7% of the time. Only the children's use of sensory/perceptual terms could predict veracity on its own, with children who talked about the games that they did not play providing more of them – a finding which may be counter-intuitive to laypersons.

In a similar study using 4- to 11-year-olds, Saykaly, Talwar, Lindsay, Bala, and Lee (2013) had children participate in a game

on three consecutive days and then asked parents to coach their children about both a game they had played and a game they had not played. The children were then interviewed on three separate occasions. Linguistic analyses were again completed, but by coders rather than using LIWC. Specifically, the researchers assessed potential differences in verbosity, cognitive operations, spontaneous corrections, admissions of lack of knowledge, temporal markers and self-references. The only differences to emerge with respect to veracity were a greater use of more spontaneous corrections in their true accounts than in their false accounts, a greater use of temporal markers in their false accounts than in their true accounts and – in the first interview only – a greater likelihood of admitting lack of knowledge in their false accounts than in their true accounts. When the pattern of language usage was assessed, a model including cognitive operations, temporal markers, self-references, admitting lack of knowledge and spontaneous corrections allowed true and false accounts to be discriminated at above-chance levels; however, greater use of cognitive operations was the only uniquely significant variable in predicting that the account was false.

One potential criticism of the aforementioned studies is the nature of the events involved: they were games rather than salient real-life experiences, and they were probably seen as moderately positive by children. Past research has clearly demonstrated that children's memory for positive events is different than their memory for more stressful events (Bauer, 2006; Peterson, 2012). A similar finding is presented by Brunet et al. (2013), who directly compared the linguistic patterns in 7- to 14-year-olds' true and false accounts of a stressful bullying event and a non-stressful sporting event; the descriptions are linguistically different. True and false accounts of the non-stressful event can be differentiated at above-chance levels using age group, word count, self-references and motion terms, but the same cannot be said for children's true and false accounts of the stressful event. The use of self-references was found to be a significant unique predictor of event veracity for the children's accounts of the non-stressful event, but the authors' initial model could not predict the veracity of the children's accounts of the stressful event. When a stepwise regression with the 29 predictors identified by Newman et al. (2003) was used, only greater use of tentative terms was shown to be a significant predictor of truthfulness when the children were describing stressful events. This finding highlights the importance of studying real-life, salient stressful events.

To date there are very few similarities amongst the variables that have been shown to predict children's veracity across studies. Age differences are consistently shown to be important (Brunet et al., 2013; Evans et al., 2012; Hauch et al., 2015; Sim & Lamb, 2013). However, when linguistic variables are considered, only the use of self-references, first-person pronouns, tentative terms (Brunet et al., 2013; Saykaly, Talwar, Lindsay, Bala, & Lee, 2013) and the discussion of affective processes (Brunet et al., 2013; Evans et al., 2012) have been identified as predictive in multiple studies. When considering the theoretical perspective, consistent with the CBCA and RM approaches, the use of perceptual processes (Evans et al., 2012) and the discussion of affective processes (Brunet et al., 2013; Evans et al., 2012) are sometimes found to be predictors of children's veracity. Finally, when comparing the findings shown with adults (Newman et al., 2003) to those shown with children, only the use of first-person pronouns (Brunet et al., 2013; Saykaly, Talwar, Lindsay, Bala, & Lee, 2013) has been identified as predictive. This suggests that linguistic analyses probably should not be used to judge children's event veracity, although so little research has been conducted that it is difficult to come to any consensus.

The differences in the predictors identified across the studies could potentially be explained by their methodological variation – for example, the differences in the events described (positive versus negative and the content of the events themselves) and the ways in which the children were coached. It may also be that differences in the bullying event versus the sporting event discussed by Brunet et al. (2013) are due to the memorability of the events themselves, as well as to the positive versus negative emotions likely evoked by participating in or talking about such events.

It has been suggested that laypersons demonstrate a greater difficulty in assessing the veracity of 4- to 9-year-olds' accounts of a stressful event than a non-stressful event (Saykaly, Talwar, Lindsay, Bala, Lee, et al., 2013). It seems important then to compare children's accounts when they are providing information about a similar stressful event. As differences are seen in the provision of information about stressful versus non-stressful events, stressful events need to be examined because they more closely parallel the events about which a child might be testifying. This is especially important given that laypersons seem to have the greatest difficulty in determining veracity when they are presented with children's true accounts of stressful events (Saykaly, Talwar, Lindsay, Bala, Lee, et al., 2013).

Another potential factor that could account for differences across studies is that the true and false stories being compared have generally been told by the same child. Saykaly, Talwar, Lindsay, Bala, Lee, et al. (2013) indicate there may be differences in the ability to differentiate between children's true versus false events based on whether the stories are being told by the same child or different children. In their study this was limited to children's telling of non-stressful accounts but it suggests that studies need to be completed where comparisons are made with different children, because when testifying, the option of having the child tell a true versus false account is generally not realistic.

Additionally, while research has been carried out with children who have been coached and with children's lies in general, to date no study using linguistic analysis has clearly compared the accounts provided by children who are telling coached versus uncoached truths or lies about a similar event using the same interview procedure. Research that has compared children who are coached to those who are not coached has generally shown that the coached children are more likely to be believed (e.g. Talwar et al., 2007; Warren et al., 2012). It seems important then to determine what it is about these children's stories that make them more believable. This is the focus of the current study – specifically, children's accounts are compared to determine whether there are differences in the various LIWC categories according to whether the children were coached or uncoached and were telling the truth or telling a lie about the stressful event of breaking a bone or being lacerated seriously enough to require suturing.

The majority of the analyses are exploratory in nature as to date no set patterns in factors have emerged as predictors of veracity. However, it was hypothesized that there would be age differences in the LIWC categories seen in children's event accounts. Consistent with Sim and Lamb's (2013) finding of differences across accounts depending on the specifics of the event and interview, it was also hypothesized that there would be differences in the LIWC variables seen and in the pattern of language used for children's coached versus uncoached accounts. Finally, it was hypothesized that the model presented by Newman et al. (2003) and the classification indicated by Sim and Lamb (which includes variables used in CBCA and RM criteria) would be found to have some utility in predicting veracity.

Method

The stories

This study uses a collection of interviews with 95 children who described the experience of an injury. All interviews were audio recorded before being transcribed verbatim. Identifying information was deleted. The children interviewed were divided into three age groups: a 'younger group' consisting of 5- to 7-year-olds ($n = 32$, $M = 5.9$, $SD = 1.0$, 16 girls), a 'middle group' consisting of 8- to 10-year-olds ($n = 32$, $M = 8.7$, $SD = 0.8$, 16 girls) and an 'older group' consisting of 11- to 14-year-olds ($n = 31$, $M = 12.5$, $SD = 1.3$, 15 girls). The children in each age group were divided into four conditions: true-coached, true-uncoached, false-coached and false-uncoached. Thus, the study has a 3 (age group) × 2 (true vs. false) × 2 (coached vs. uncoached) design with 8 children in each cell, but with the exception that there is one less 11- to 14-year-old girl in the uncoached true condition.

The accounts for the true-uncoached condition 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 (see Peterson, 2010). Additional children who had been injured were recruited and coached by their parents prior to being interviewed in order to form the true-coached condition. All true accounts were corroborated by a parent. The fabricated accounts were prepared by children who had never had a comparable injury (this was verified by their parents). The children in this condition were recruited either through a childcare center or through friends' and students' connections. The children in the false-uncoached condition were asked to fabricate a story about an injury that was similar to those sustained by the injured children. For the children in the false-coached condition, the parents were given a list of questions covering the types of details that a child might be expected to know and were asked to help their child practice her or his story at least once per day for the four days leading up to the interview, for approximately 10 to 20 minutes per session. The 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 and 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 a statement. This interview has been used in prior research (e.g. Peterson, 1999; Peterson & Bell, 1996; Peterson & Whalen, 2001). The same interview procedure was used for all of the children, with the exception that those in the false-account conditions were told at the beginning of the interview to 'try to fool other people' into believing that the injury had really happened to them. These transcripts have been used in previous studies comparing the ability of laypersons to determine the veracity of children's truthful versus deceitful coached and uncoached accounts (Warren et al., 2012, 2015).

LIWC analysis

Each transcribed story was entered into a text file and prepared for LIWC analysis. Interviewer questions were removed, as were misspellings, abbreviations and some contractions (those identified as problematic by LIWC). LIWC 2007 can complete analysis for 72 LIWC variables (Pennebaker, Chung, Ireland, Gonzales, & Booth, 2007). However, only the 29 variables used in Newman et al. (2003) were selected for the present statistical analyses. No additional variables were removed, as no extremely low rates of usage (less than 0.2% of the time) were found for any of these 29 variables.

Results

Analyses were carried out to assess differences in the provision of specific LIWC categories according to age and gender, presence of coaching and event veracity. Additionally, the specific pattern of language use was assessed. The initial analyses revealed no main effects of gender and no interactions between gender and the other variables that were assessed. Consequently, gender is not included in the analyses presented below.

Differences in the provision of LIWC categories

In order to determine whether or not there are differences in the use of the various LIWC categories, 3 (5–7 years, 8–10 years, 11–14 years) × 2 (coached vs. uncoached) × 2 (true vs. false) between-subjects analyses of variance (ANOVAs) were completed for each of the 29 LIWC categories identified by Newman et al. (2003) as being of potential importance. As the analyses are mainly exploratory in nature, no correction factor is used for alpha and differences with an alpha level of less than .05 are considered significant. With the large number of comparisons and no correction factor, differences should be considered with caution. The findings are presented according to the variable of interest (age, presence of coaching and event veracity).

Age

Age differences in the children's stories were found with respect to word count, $F(2, 83) = 7.95$, $p = .001$, partial $\eta^2 = .16$, the overall use of words pertaining to affect, $F(2, 83) = 4.23$, $p = .018$, partial $\eta^2 = .09$, the discussion of positive emotions, $F(2, 83) = 6.16$, $p = .003$, partial $\eta^2 = .13$, the use of cognitive mechanisms, $F(2, 83) = 4.53$, $p = .014$, partial $\eta^2 = .10$, and the use of motion terms, $F(2, 83) = 3.60$, $p = .032$, partial $\eta^2 = .08$. Pairwise comparisons were completed to determine the specific breakdown of age differences; it was found that the 11- to 14-year-olds provided significantly higher word counts ($M = 414.36$, $SD = 221.31$) than both the 5- to 7-year-olds ($M = 249.97$, $SD = 135.56$) and the 8- to 10-year-olds ($M = 290.31$, $SD = 166.26$), mean difference = 164.39, $p < .001$, 95% CI [79.08, 249.70] and mean difference = 124.05, $p = .005$, 95% CI [38.75, 209.35], respectively. There is no difference in the word counts provided by the 5- to 7-year-olds and the 8- to 10-year-olds.

When overall discussion of affect is considered, the 11- to 14-year-olds ($M = 3.43$, $SD = 1.66$) provided significantly more words describing emotions than the 5- to 7-year-olds ($M = 2.38$, $SD = 1.28$), mean difference = 1.06, $p = .006$, 95% CI [0.32, 1.80]. No other age differences were found. Pairwise comparisons assessing the provision of positive emotions show that both the 8- to 10-year-olds ($M = 1.75$, $SD = 1.40$) and the 11- to 14-year-olds ($M = 1.90$, $SD = 1.34$) provided significantly more positive emotions than the 5- to 7-year-olds ($M = 0.93$, $SD = 0.92$), mean difference = 0.82, $p = .007$, 95% CI [0.23, 1.40] and mean difference = 0.97, $p = .002$, 95% CI [0.38, 1.56], respectively. Finally, the 11- to 14-year-olds provided significantly more cognitive mechanisms ($M = 17.03$, $SD = 3.70$) than the 5- to 7-year-olds ($M = 14.23$, $SD = 3.44$), mean difference = 2.80, $p = .004$, 95% CI [0.94, 4.66], and the 11- to 14-year-olds ($M = 4.44$, $SD = 1.50$) provided significantly fewer motion terms than the 5- to 7-year-olds ($M = 5.46$, $SD = 1.23$), mean difference = −1.01, $p = .009$, 95% CI [− 1.77, −0.26].

In addition to the main effects of age, there are also a number of interactions between age, presence of coaching and event veracity. Specifically, there are significant age × presence of coaching interactions for the provision of insight terms, $F(2, 83) = 3.78$, $p = .027$, partial $\eta^2 = .08$, and the use of the present tense, $F(2, 83) = 3.68$, $p = .029$, partial $\eta^2 = .08$. There is no age difference in the provision of insight terms for the

children who were coached, $F(2, 45) = 1.58$, $p = .217$, partial $\eta^2 = .07$, but there is an age difference for the children who were uncoached, $F(2, 44) = 3.87$, $p = .028$, partial $\eta^2 = .15$. In particular, the uncoached 8- to 10-year-olds ($M = 1.16$, $SD = 0.96$) provided significantly more insight words than both the 5- to 7-year-olds ($M = 0.47$, $SD = 0.66$), mean difference = 6.93, $p = .013$, 95% CI [0.15, 1.23], and the 11- to 14-year-olds ($M = 0.57$, $SD = 0.59$), mean difference = 0.59, $p = .035$, 95% CI [0.04, 1.14]. Pairwise comparisons indicate no significant age differences for coached versus uncoached children in the use of the present tense; however, both the 5- to 7-year-olds and the 8- to 10-year-olds used the present tense more frequently in the uncoached condition than in the coached condition, whereas the opposite is true for the 11- to 14-year-olds.

There is also an age × event veracity interaction for the provision of positive emotions, $F(2, 83) = 4.28$, $p = .017$, partial $\eta^2 = .09$. There is a difference in the provision of positive emotions for true versus false accounts for just the 8- to 10-year-olds, $F(1, 30) = 12.66$, $p = .001$, partial $\eta^2 = .30$, 95% CI [0.66, 2.43], wherein those who were telling a lie ($M = 2.52$, $SD = 1.56$) provided more positive emotion words than those who were telling the truth ($M = 0.98$, $SD = 0.75$).

Presence of coaching

Differences according to whether the children were coached or uncoached were found with respect to the word count, $F(1, 83) = 8.90$, $p = .004$, partial $\eta^2 = .10$, and the use of dictionary words, $F(1, 83) = 6.15$, $p = .015$, partial $\eta^2 = .07$. The children who were coached ($M = 370.29$, $SD = 203.92$) provided higher words counts than the children who were uncoached ($M = 266.14$, $SD = 157.25$), mean difference = 104.16, $p = .004$, 95% CI [34.71, 173.60]. In contrast to this, the children who were uncoached ($M = 92.14$, $SD = 5.19$) provided more dictionary words than the children who were coached ($M = 89.99$, $SD = 2.77$), mean difference = 2.14, $p = .015$, 95% CI [0.43, 3.86].

In addition to these main effects, there are a number of presence of coaching × event veracity interactions, specifically for the use of prepositions, $F(1, 83) = 6.18$, $p = .015$, partial $\eta^2 = .07$, the discussion of negative emotions, $F(1, 83) = 4.28$, $p = .010$, partial $\eta^2 = .08$, the use of sensory or perceptual terms, $F(1, 83) = 5.00$, $p = .028$, partial $\eta^2 = .06$, the discussion of space, $F(1, 83) = 6.45$, $p = .013$, partial $\eta^2 = .07$, the use of inclusive terms, $F(1, 83) = 9.26$, $p = .003$, partial $\eta^2 = .10$, and the use of the present tense, $F(1, 83) = 8.89$, $p = .004$, partial $\eta^2 = .10$.

Differences in the use of prepositions, $F(1, 45) = 16.35$, $p < .001$, partial $\eta^2 = .27$, 95% CI [1.17, 3.48], sensory or perceptual terms, $F(1, 45) = 7.05$, $p = .011$, partial $\eta^2 = .14$, 95% CI [0.20, 1.44], space terms, $F(1, 45) = 15.07$, $p < .003$, partial $\eta^2 = .25$, 95% CI [1.15, 3.62], inclusive terms, $F(1, 45) = 5.23$, $p = .027$, partial $\eta^2 = .10$, 95% CI [0.25, 3.95], and the present tense, $F(1, 45) = 4.52$, $p = .039$, partial $\eta^2 = .09$, 95% CI [−2.24, −0.60], were found – but only for the uncoached children. It was discovered that the uncoached children who were telling the truth used more prepositions ($M = 12.98$, $SD = 2.21$ vs. $M = 10.66$, $SD = 1.71$), sensory or perceptual terms ($M = 2.09$, $SD = 1.22$ vs. $M = 1.23$, $SD = 0.86$), space terms ($M = 9.14$, $SD = 2.38$ vs. $M = 6.76$, $SD = 1.80$), and inclusive terms ($M = 9.86$, $SD = 3.51$ vs. $M = 7.76$, $SD = 2.76$) than the uncoached children who were telling a lie. Somewhat contradictory to this, the uncoached children who were telling a lie ($M = 3.95$, $SD = 2.05$) used the present tense more often than the uncoached children who were telling the truth ($M = 2.80$, $SD = 1.62$).

When the use of negative emotions was examined, in contrast to the other presence of coaching × event veracity interactions, a difference was found for just the children who were coached, $F(1, 46) = 5.13$, $p = .028$, partial $\eta^2 = .10$, 95% CI [0.05, 0.87] –

specifically, the coached children who were telling a lie ($M = 1.68$, $SD = 0.63$) provided more negative emotions than the coached children who were telling the truth ($M = 1.22$, $SD = 0.77$).

Event veracity

Differences according to whether children were telling the truth or a lie were found with respect to the use of articles, F (1, 83) = 5.29, $p = .024$, partial $\eta^2 = .06$, 95% CI [−1.79, −0.13], the use of prepositions, F (1, 83) = 7.62, $p = .007$, partial $\eta^2 = .08$, 95% CI [0.35, 2.12], the use of positive emotions, F (1, 83) = 5.09, $p = .027$, partial $\eta^2 = .06$, 95% CI [−1.03, −0.07], and discussion of space, F (1, 83) = 7.83, $p = .006$, partial $\eta^2 = .09$, 95% CI [0.36, 2.13]. The children who were telling the truth used fewer articles ($M = 6.24$, $SD = 1.65$ vs. $M = 7.20$, $SD = 2.21$) and fewer positive emotions ($M = 1.25$, $SD = 1.08$ vs. $M = 1.80$, $SD = 1.46$) than the children who were telling a lie. However, the findings considering the use of articles need to be viewed with caution, as the aforementioned presence of coaching × event veracity interaction suggests that this is only the case for the uncoached children. It was also found that the children who were telling the truth used more prepositions ($M = 12.30$, $SD = 2.34$ vs. $M = 11.06$, $SD = 2.08$) and more terms to describe space ($M = 8.33$, $SD = 2.32$ vs. $M = 7.08$, $SD = 2.11$) than the children who were telling a lie. The finding considering the use of terms to describe space should also be viewed with caution, as the presence of coaching × event veracity interaction again suggests that this is only the case for the uncoached children.

The pattern of language used

As noted in the introduction, it has been argued that just looking at individual LIWC cues may not capture the complete picture that could indicate differences between children's true and false accounts (Newman et al., 2003). Thus, binary logistic regression analyses were completed in order to determine whether or not there are differences in the linguistic patterns seen for children's true versus false stories. Consistent with Sim and Lamb's (2013) findings, in the present study there are age differences in the children's use of some of the LIWC categories that are of interest. To control for these differences, age was forced into the following regression models prior to any of the LIWC categories. First, overall regression equations were completed using all children. Additionally, as presence of coaching is a variable of interest and differences in the use of the LIWC categories across the coaching conditions have been identified, children's coached and uncoached accounts were examined separately.

Overall regression. Initially, a forward stepwise (likelihood ratio) binary logistic regression was completed using veracity as the outcome variable and age and the 29 LIWC variables of interest as potential predictor variables. When veracity is predicted from age alone, the model is not significant, Model $\chi^2(1) = 0.02$, $p = .90$, $R^2_N = .00$. However, two potential models for predicting veracity emerged. The model that includes age and the use of spatial terms is significant, Model $\chi^2(2) = 7.17$, $p = .028$, $R^2_N = .10$. The prediction accuracy improves from 50.5% (-2LL = 131.67) using just the constant to 65.3% (61.7% for truth-tellers, 68.8% for liars) with the two predictors added (-2LL = 124.52). The analysis revealed that the use of spatial terms is a significant predictor of veracity, Wald = 6.49, $p = .011$. As shown in Step 2 in Table 1, the odds ratio (OR) associated with spatial terms suggests that as the use of spatial terms increases, it becomes more likely that the child is telling the truth. The model that includes age, the use of spatial terms and the use of articles is also significant, Model $\chi^2(3) = 17.90$, $p < .001$, $R^2_N = .23$. The prediction accuracy improves to 69.5% (63.8% for truth-tellers, 75.0% for liars) with the three predictors added (-2LL =

Table 1. The logistic regression model predicting event veracity for all children.

			95% CI for odds ratio		
		B (*SE*)	Lower	Odds ratio	Upper
Step 1					
	Constant	−0.04 (0.54)			
	Age group	0.03 (0.25)	0.63	1.03	1.69
Step 2					
	Constant	2.02 (0.99)			
	Age group	−0.03 (0.26)	0.58	0.98	1.63
	Space	−0.25* (0.10)	0.64	0.78	0.94
Step 3					
	Constant	0.12 (1.17)			
	Age group	−0.01 (0.28)	0.57	0.99	1.70
	Space	−0.38** (0.12)	0.54	0.69	0.87
	Articles	0.42** (0.14)	1.15	1.53	2.02

Note: $^{*}p < .05$; $^{**}p < .01$.

113.79). Both the use of spatial terms, Wald = 10.12, $p = .001$, and the use of articles, Wald = 8.79, $p = .003$, are significant predictors of veracity. As with the first model, the odds ratio suggests that as the use of spatial terms increases, it becomes more likely that the child is telling the truth. In contrast, the odds ratio for the use of articles suggests that as usage increases, it becomes more likely that the child is telling a lie.

Next, a binary logistic regression analysis was completed using veracity as the outcome variable and using the enter method to force age and the variables previously identified by Newman et al. (2003) (first-person singular pronouns, third-person pronouns, negative emotion words, exclusive words and motion words) into the model. The overall model is not significant and no variables emerged as unique predictors of veracity.

Finally, a binary logistic regression analysis was completed using veracity as the outcome variable and using the enter method to force age and the LIWC categories that correspond with CBCA and RM criteria (affective processes, perceptual processes, feel words, space words, time words and insight words) into the model. The overall model is significant, Model $\chi^2(7) = 14.56$, $p = .042$, $R^2_N = .19$, and the prediction accuracy improves from 50.5% (-2LL = 131.69) using just the constant to 67.4% (66.0% for truth-tellers, 68.8% for liars) with the variables added (-2LL = 117.13). None of the individual variables are significant unique predictors of veracity.

Uncoached accounts. When a forward stepwise (likelihood ratio) binary logistic regression was completed with veracity as the outcome variable and age and the 29 LIWC categories that are of interest as potential predictor variables, the result was a complete statistical separation of the data (it appeared to indicate that 100% of the variability was accounted for). This may well be due to the low number of stimuli (35 accounts) and the high number of predictor variables (30). Potential predictors of veracity for children's uncoached accounts include the use of negations, prepositions, negative emotions, sensory or perceptual terms, spatial terms and the past tense. As suggested by Heinze and Schemper (2002), in order to correct for this separation, additional analyses were completed using Firth binary logistic regression.

With the Firth correction factor, no combined model or individual variables emerged as predictors of veracity.

A binary logistic regression analysis (enter method) was completed with veracity as the outcome variable and age and the variables identified by Newman et al. (2003) (first-person singular pronouns, third-person pronouns, negative emotion words, exclusive words and motion words) as predictor variables. The overall model is not significant and no variables emerged as potential predictors of veracity.

Finally, a binary logistic regression analysis (enter method) was completed with veracity as the outcome variable and age and the LIWC categories that correspond with CBCA and RM criteria (affective processes, perceptual processes, feel words, space words, time words and insight words) as predictor variables. The overall model is significant, Model $\chi^2(7) = 19.32, p = .007, R^2_N = .45$. The prediction accuracy improves from 51.1% (-2LL = 65.14) using just the constant to 80.9% (78.3% for truth-tellers, 83.3% for liars) with the predictor variables added (-2LL = 45.81). The use of spatial terms was the only unique predictor of veracity, Wald = 5.92, $p = .015$. The odds ratio, 0.56, 95% CI [0.36, 0.89], suggests that as the use of spatial terms increases, it becomes more likely that the child is telling the truth.

Coached accounts. As with the overall analysis and the analysis of uncoached accounts, initially a forward stepwise binary logistic regression (likelihood ratio) was completed with veracity as the outcome variable and age and the 29 LIWC categories as predictor variables. When the veracity of the coached accounts is predicted from age alone, the model is not significant, Model $\chi^2(1) = 0.00$, $p = 1.00, R^2_N = .00$. However, two potential models emerged for predicting veracity. The model that includes age and the use of negations is significant, Model $\chi^2(2) = 7.36, p = .025, R^2_N = .19$. The prediction accuracy improves from 50.0% (-2LL = 66.54) using just the constant to 64.6% (54.2% for truth-tellers, 75.0% for liars) with the two predictors added (-2LL = 59.18). The analysis reveals that the use of negations is a significant predictor of veracity, Wald = 5.80, $p = .016$. As shown in Step 2 in Table 2, the odds

Table 2. The logistic regression model predicting event veracity for coached children.

			95% CI for odds ratio		
		B (*SE*)	Lower	Odds ratio	Upper
Step 1					
	Constant	−0.00 (0.76)			
	Age group	0.00 (0.35)	0.50	1.00	2.00
Step 2					
	Constant	0.88 (0.90)			
	Age group	0.16 (0.39)	0.55	1.17	2.51
	Negations	−0.77* (0.32)	0.25	0.46	0.87
Step 3					
	Constant	−0.49 (1.11)			
	Age group	0.07 (0.41)	0.48	1.08	2.40
	Negations	−0.86* (0.35)	0.21	0.43	0.85
	Negative emotions	1.16* (0.55)	1.08	3.18	9.31

Note: *$p < .05$.

ratio associated with negations suggests that as the use of negations increases, it becomes more likely that the child is telling the truth. The model that includes age, the use of negations and the use of negative emotions is also significant, Model $\chi^2(3) = 12.82$, $p = .005$, $R^2_N = .31$. The prediction accuracy improves to 70.8% (66.7% for truth-tellers, 75.0% for liars) with the three predictors added (-2LL = 53.72). Both the use of negations, Wald = 5.87, $p = .015$, and the use of negative emotions, Wald = 4.43, $p = .035$, are significant predictors of veracity. As with the first model, the odds ratio suggests that as the use of negations increases, it becomes more likely that the child is telling the truth. In contrast, the odds ratio for the discussion of negative emotions suggests that as usage increases, it becomes more likely that the child is telling a lie.

A binary logistic regression analysis (enter method) was then completed with veracity as the outcome variable and age and the variables identified by Newman et al. (2003) (first-person singular pronouns, third-person pronouns, negative emotion words, exclusive words and motion words) as predictor variables. The overall model is significant, Model $\chi^2(6) = 13.44$, $p = .037$, $R^2_N = .33$. The prediction accuracy improves from 50.0% (-2LL = 66.54) using just the constant to 68.8% (66.7% for truth-tellers, 70.8% for liars) with the predictor variables included (-2LL = 53.11). The use of negative emotions is the only unique predictor of veracity, Wald = 6.26, $p = .012$. The odds ratio, 4.05, 95% CI [1.35, 12.10], suggests that as the use of negative emotions increases, it becomes more likely that the child is telling a lie.

Finally, a binary logistic regression analysis (enter method) was completed with age and the LIWC categories that correspond with CBCA and RM criteria (affective processes, perceptual processes, feel words, space words, time words and insight words). The overall model is not significant and no variables emerged as potential predictors of veracity.

Discussion

This study was undertaken in order to search for linguistic differences in children's uncoached versus coached true and false accounts of the stressful experience of sustaining an injury. Because past research has indicated a difficulty on the part of laypersons in determining differences between children's true versus false accounts, it is important to determine whether or not linguistic software can be used to differentiate between such accounts (e.g. Talwar et al., 2007; Vrij et al., 2002, 2004). This has been used with some success in evaluating adults' event descriptions (Hauch et al., 2015), but to date very little research has examined the linguistic content of children's event descriptions (Brunet et al., 2013; Evans et al., 2012; Saykaly, Talwar, Lindsay, Bala, & Lee, 2013). Furthermore, it has been suggested that the use of linguistic software may be inappropriate for examining the veracity of children's accounts (Sim & Lamb, 2013). However, to date, with such a limited amount of research on the topic, potential successes and failures with the use of such software are unclear. Some of the problems with using linguistic software to evaluate children's accounts are illustrated in the present study.

An issue which has been mentioned in the past is that there are age differences in the provision of information which may limit the generalizability of models developed in the adult literature to the evaluation of children's accounts (Sim & Lamb, 2013). Supporting this, a number of studies clearly indicate age differences in children's provision of event details (Brunet et al., 2013; Evans et al., 2012; Sim & Lamb, 2013). It was thus hypothesized that age differences would emerge in the children's accounts in the present study – and the younger children did indeed provide accounts that are linguistically different compared to those of the older children. In particular, there are differences in linguistic categories including the word count, the use of cognitive mechanisms and

the use of affect. These differences point to changes that children are going through developmentally, which – as indicated by Sim and Lamb (2013) – suggests that it may be inappropriate to use models developed to assess adults' accounts to similarly judge the veracity of children's accounts.

In addition to its main effects, age also interacted with presence of coaching for the use of insight terms and positive emotions. Sim and Lamb (2013) argue that there are differences in the linguistic breakdown of event accounts across a number of variables of interest (e.g. the style in which the interview is conducted). This led to the hypothesis in the present study that coaching could potentially affect the linguistic breakdown of children's stories. Supporting this, with respect to the use of both insight terms and positive emotions, it seems as though – at least for the 5- to 10-year-olds – the coaching changed the pattern of language that was used in a way that made the lies more similar to the truth.

Differences in coached versus uncoached accounts continued to emerge in direct comparisons. For example, the coached children told longer stories. Intuitively, one might assume that lengthier stories indicate true accounts of actual events because children are able to provide more information. If children's coached accounts are longer, it seems fair to assume that this leads to difficulties on the part of laypersons who are judging the veracity of children's coached accounts. On the surface, uncoached children's greater use of dictionary words might appear to undermine the notion of coaching leading to a more believable story. However, it seems plausible that children are coached to provide simple and easy to understand stories that the adults themselves would consider believable. Such a practice might also remove the need for children to use their already limited working memory to remember complex terminology, thus allowing them to better focus on their story and make it more believable.

As a whole, a number of implications arise from the present findings in relation to why coaching can be successful in allowing children to tell a more believable story. This is particularly evident when the interactions between event veracity and presence of coaching are considered. In the majority of cases, the differences between the children's true versus false accounts exist when the children were uncoached, with the exception of the use of negative emotions, which was more likely when the children were coached.

These differences emphasize the importance of considering coached accounts differently from uncoached accounts. In the eyewitness literature, it has been seen that jurors have difficulty believing witnesses who show little or no negative emotion (e.g. Wessel, Magnussen, & Melinder, 2013). Laypersons then appear to believe that a truthful story is one which contains negative affect, potentially explaining why an adult would coach a child in such a manner. These findings suggest that the use of linguistic software might help in determining the veracity of a story, but only if it is known whether or not the child telling it has been coached – a possibility that is very unlikely in real-world evaluations.

Because previous studies indicate the importance of assessing the whole story and not its individual parts (e.g. Newman et al., 2003), regression analyses were used to evaluate the true versus false accounts. As noted in the introduction, at present there is no clear identification of the predictors of veracity – and so no hypotheses were formed. When stepwise analyses were used in an attempt to determine possible predictors of veracity, despite a somewhat impressive ability to determine event veracity from linguistic analysis, differences emerged as a consequence of whether the data were analyzed overall or for uncoached versus coached accounts. When the analysis was completed for all accounts, a model that includes the use of spatial terms and articles emerged, with both as unique predictors. The use of articles

has not been identified as a predictor of veracity in any of the past research assessing linguistic differences in children's true versus false accounts, and although the use of spatial terms is suggested as a potential predictor in the CBCA and RM approaches (Sim & Lamb, 2013) it has not previously been seen as a unique predictor of veracity in any of the other studies assessing children's veracity from a linguistic perspective (Brunet et al., 2013; Evans et al., 2012; Saykaly, Talwar, Lindsay, Bala, & Lee, 2013).

In relation to the uncoached accounts, although the statistical analysis is somewhat questionable, possible predictors include the use of negations, prepositions, negative emotions, sensory or perceptual terms, spatial terms and the past tense. Negations, prepositions and past tense are not proposed as potential indicators of veracity in the CBCA and RM approaches, and only the use of past tense is somewhat supported by past research (Saykaly, Talwar, Lindsay, Bala, & Lee, 2013). The use of spatial terms is suggested as a potential predictor in the CBCA and RM approaches, but has not been identified in past research. Both the use of negative emotions and the use of sensory perceptual terms are also a part of the CBCA and RM approaches, and both have been identified in at least some previous research.

In relation to the coached accounts, a model that includes the use of negations and negative emotions emerged, and both are unique predictors of veracity. Again, as noted, the use of negative emotions has been suggested as a possible predictor in both Newman et al. (2003) and the CBCA and RM approaches, as well as being identified in past research. Furthermore, it is in the direction that has been suggested (Hauch et al., 2015), i.e. that liars might be expected to provide more negative emotions, perhaps because they feel more negative emotions when lying. However, the use of negations has not been identified as a potential predictor in past research assessing children's veracity.

In addition to examining potential predictors of veracity, it is important to examine the applicability of the variables identified in past research. More specifically, Newman et al. (2003) identify possible predictors of adult veracity that could potentially be used to determine children's veracity, but this has not really been assessed to date. In the present study, when the overall model suggested by Newman et al. was examined, in contrast to the hypothesis, when all of the children's accounts were assessed and when just the uncoached accounts were assessed, the overall model was not found to be significant. In contrast to this, the overall model and the unique predictor of negative emotions were found to be significant when the children's coached accounts were examined. This supports the suggestion by Sim and Lamb (2013) that it is important to be cognizant of the differences between children and adults rather than assuming that models identified in the adult literature can be applied to children.

In contrast to the model identified by Newman et al. (2003), the CBCA (Raskin & Esplin, 1991) and RM (Johnson & Raye, 1981) approaches are seen as being applicable to children. It has been argued that there are age differences in the use of some of the criteria used (Buck, Warren, Betman, & Brigham, 2002; Vrij et al., 2002), but the variables themselves apply to children. Somewhat supporting this and the hypothesis of the present study, the overall model using the CBCA and RM criteria was found to be significant when all of the children's accounts were analyzed, as well as in the analysis of the children's uncoached accounts. In addition, although there are no unique predictors in the analysis that includes all of the children, the use of spatial terms emerged as a significant predictor in assessing uncoached accounts. Further demonstrating the problem with such approaches, however, the model was not found to be a significant predictor in evaluating the children's coached accounts. Again, as it is unlikely in practice that evaluators will know whether they are evaluating an

uncoached or a coached account, this makes the use of such an analysis questionable in a real-world setting.

The present study is not without limitations. Perhaps most importantly, the small sample size makes it difficult to evaluate the children's event accounts using regression analysis. Future research should include a larger number of stimuli in order to resolve this problem. Furthermore, the parents were given instructions on how to coach their children and no formal measure of coaching was used, thus limiting the generalizability to actual coached accounts and making it difficult to determine what facets of coaching led to the formulation of such believable stories.

An interesting finding that future researchers might want to assess however is that the variables suggested by Newman et al. (2003) were found to be significant for the coached but not uncoached accounts. This leads to the question of whether or not coached accounts are more similar to adults' accounts than to other children's uncoached accounts when making evaluations of event veracity, a possibility that is worth exploring.

When the findings of the present study are compared to the existing literature, very few similarities are seen between the variables that emerged as predictors of veracity here compared to other studies. The exception is the finding that the use of negative emotions, the use of sensory or perceptual processes and the use of spatial terms are potentially predictive when children provide uncoached accounts, which is consistent with the CBCA and RM approaches and with some past research. The failure to find many similarities with past research and current theory is not entirely surprising however, given the lack of similarities seen in the review of past research presented herein.

Overall, what is clear is that these findings suggest that linguistic analysis – at least using the variables which have been suggested to date – cannot clearly predict event veracity. In fact, as suggested by Sim and Lamb (2013), with so many other variables influencing the content of children's event accounts, it may be inappropriate to use such analyses to evaluate their veracity. In addition, there are differences in the ways in which children provide true versus false accounts depending on their age and on whether or not they have been coached. This suggests a reason why laypersons have difficulty differentiating between children's true and false accounts – especially when coaching is involved. Crucially, the differences that were found between the true and false accounts of the uncoached children disappeared when the children were coached. This means that without knowing whether or not a child has been coached, conducting a linguistic analysis of her or his account will not help to judge its veracity. The coached accounts of liars and truth-tellers could not be differentiated between in this study – and in the real world, information about children's coaching history is seldom known.

Acknowledgments

We would like to thank all the recruiters, interviewers and transcribers who participated in this research, and especially Penny Voutier for her data management and Daniel Nadolny for statistical consultation. Most of all we would like to thank the parents and children who allowed us into their homes and who cooperated in the completion of this project.

Funding

This work was supported by the Natural Sciences and Engineering Research Council of Canada [grant number 513-2012]. Additional funding came from the Memorial University Undergraduate Career Experience Program and the Student Work and Service Program.

Ethical standards

Declaration of conflicts of interest

Kelly L. Warren has declared no conflicts of interest.

Carole Peterson has declared no conflicts of interest.

Cassy C. Gillingham has declared no conflicts of interest.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee (TCPS-2) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

Informed consent was obtained from all individual participants included in the study.

References

Ali, M., & Levine, T. (2008). The language of truthful and deceptive denials and confessions. *Communication Reports, 21*, 82–91. doi: 10.1080/08934210802381862

Bala, N., Lee, K., & McNamara, E. (2001). Children as witnesses: Understanding their capacities, needs, and experiences. *Journal of Social Distress and the Homeless, 10*, 41–68. doi: 10.1023/A:1009429602266

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: John Wiley & Sons Inc.

Bond, G. D., & Lee, A. Y. (2005). Language of lies in prison: Linguistic classification of prisoners' truthful and deceptive natural language. *Applied Cognitive Psychology, 19*, 313–329. doi: 10.1002/acp.1087

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, S. M., Warren, A. R., Betman, S., & Brigham, J. C. (2002). Age differences in Criteria-Based Content Analysis scores in typical child sexual abuse interviews. *Journal of 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

Evans, A. D., Brunet, M. K., Talwar, V., Bala, N., Lindsay, R. C. L., & Lee, K. (2012). The effects of repetition and children's true and false reports. *Psychiatry, Psychology, and Law, 19*, 517–529. doi: 10.1080/13218711119.2011.615808

Faller, K. C. (2007). Coaching children about sexual abuse: A pilot study of professionals' perceptions. *Child Abuse and Neglect, 31*, 947–959. doi:10.1016/j.chiabu.2007.005.004

Hauch, V., Blandon-Gitlin, I., Masip, J., & Sporer, S. L. (2015). Are computers effective lie detectors? A meta-analysis of linguistic cues to deception. *Personality and Social Psychology Review, 19*, 307–342. doi: 10.1177/1088868314556539

Heinze, G., & Schemper, M. (2002). A solution to the problem of separation in logistic regression. *Statistics in Medicine, 21*, 240–2419. doi: 10.1002/sim.1047

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67–85. doi: 10.1037/0033-295X.88.1.67

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

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin, 29*, 665–675. doi: 10.1177/0146167203251529

Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2007). *The development and psychometric properties of LIWC2007*. Austin, TX: LIWC.net.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic inquiry and word count (LIWC): LIWC 2001*. Mahwah, NJ: Lawrence Erlbaum.

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

Raskin, D. C., & Esplin, P. W. (1991). Statement validity assessment: Interview procedures and content analysis of children's statements of sexual abuse. *Behavioral Assessment, 12*, 265–291.

Saykaly, C., Talwar, V., Lindsay, R. C. L., Bala, N. C., & Lee, K. (2013). The influence of multiple interviews on the verbal markers of children's deception. *Law and Human Behavior, 37*, 187–196. doi: 10.1037/lhb0000023

Saykaly, C., Talwar, V., Lindsay, R. C. L., Bala, N. C., Lee, K., Bertrand, M., & 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

Sim, M. P. Y., & Lamb, M. E. (2013). Children's disclosure of sexual abuse: How motivational factors affect linguistic categories related to deception detection. *Psychology, Crime, and Law, 19*, 649–660. doi: 10.1080/1068316X.2012.719621

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. (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. (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., 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. (2007). White lie-telling in children for politeness purposes. *International Journal of Behavioral Development, 31*, 1–11. doi:10.117/0165025406073530

Trocme, N., & Bala, N. (2005). False allegations of abuse and neglect when parents separate. *Child Abuse and Neglect, 29*, 1333–1345. doi: 10.106/j.chiabu.2004.06.016

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

Warren, K. L., Bkhtiar, A., Mulrooney, B., Raynor, G., Dodd, E., & Peterson, C. (2015). Adults detection of deception in children: Effect of coaching and age for children's true and fabricated reports of injuries. *Behavioral Sciences and the Law, 33*, 784–800. doi: 10.1002/bsl.2210

Warren, K. L., Dodd, E., Raynor, G., & Peterson, C. (2012). Detecting children's lies: Comparing true accounts about highly stressful injuries with unprepared, prepared, and coached lies. *Behavioral Sciences and the Law, 30*, 329–341. doi: 10.1002/bsl.1994

Wessel, E., Magnussen, S., & Melinder, A. M. D. (2013). Expressed emotions and perceived credibility of child mock victims disclosing physical abuse. *Applied Cognitive Psychology, 27*, 611–616. doi: 10.1002/acp.2935