How Children Report True and Fabricated Stressful and Non-Stressful Events[1]

Megan K. Brunet[a], **Angela D. Evans**[b], **Victoria Talwar**[c], **Nicholas Bala**[d], **Rod C. L. Lindsay**[d], and **Kang Lee**[a,*]

[a]University of Toronto, Canada

[b]Brock University, Canada

[c]McGill University, Canada

[d]Queen's University, Canada

Abstract

As children can be victims or witnesses to crimes and may be required to testify about their experiences in court, the ability to differentiate between children's true and fabricated accounts of victimization is an important issue. This study used automated linguistic analysis software to detect linguistic patterns in order to differentiate between children's true and false stressful bullying reports and reports of non-stressful events. Results revealed that children displayed different linguistic patterns when reporting true and false stressful and non-stressful stories, with non-stressful stories being more accurately discriminated based on linguistic patterns. Results suggest that it is difficult to discriminate accurately and consistently between children's true and false stories of victimization.

Keywords

child witness testimony; children; deception; linguistic analysis

Introduction

The assessment of children's credibility is a challenge faced by various professions such as social workers, psychologists, police officers, lawyers and judges, who must regularly decide how to respond to children's reports of victimization. Many professionals are concerned that children may not be capable of making accurate accounts, or that children may fabricate false reports of victimization to gain some advantage or to satisfy authority figures (Bala, Ramakrishnan, Lindsay, & Lee, 2005). While some studies have revealed that children can give highly accurate accounts and can make competent witnesses (Bruck & Ceci, 1999; Quas, Goodman, Ghetti, & Redlich, 2000), other studies indicate that children can be coached into telling convincing fabricated accounts of events, or "deliberate lies" (Lyon, Malloy, Quas, & Talwar, 2008; Orcutt, Goodman, Tobey, Batterman-Faunce, & Thomas, 2001; Talwar, Lee, Bala, & Lindsay, 2006). As such, further research is needed to

---

Correspondence: Kang Lee, OISE/University of Toronto, 45 Walmer Rd, Toronto, Ontario M5R 2X2, Canada. Tel: +1 416 934 4597. Fax: +1 416 934 4565. kang.lee@utoronto.ca.

establish methods to accurately distinguish between children's truthful and fabricated statements.

Extensive research has revealed that after committing a minor misdeed (e.g., peeking at a forbidden toy), children as young as three can and will lie in an effort to conceal their own transgressions (Lewis, 1993; Lewis, Stanger, & Sullivan, 1989; Peskin, 1992; Talwar, Gordon, & Lee, 2007; Talwar & Lee, 2002; Talwar, Lee, Bala, & Lindsay, 2002; Talwar, Lee, Bala, & Lindsay, 2004; for review see Talwar & Lee, 2008). Furthermore, some children will lie to conceal another person's transgression, especially if that person is a parent or authority figure (Talwar et al., 2006; Tye, Amato, Honts, Devitt, & Peters, 1999).

However, most studies considering children's lie-telling abilities refer to a minor misdeed or a fabricated account of neutral or pleasant events that may not be ecologically relevant or applicable to legal settings (Blandon-Gitlin, Pezdek, Rogers, & Brodie, 2005; Newcombe & Bransgrove, 2007; Strömwall, Granhag, & Landström, 2007; Talwar & Lee, 2002; Talwar et al., 2006; Tye et al., 1999; Vrij, Akehurst, Soukara, & Bull, 2004a). When children's credibility is called into question outside the laboratory setting, the situation is often serious and the topic can be stressful or emotional to discuss. As such, it is important to consider how children's reports differ for both stressful and non-stressful events. Research considering children's reports of natural disasters reveal that when moderately stressed, children will remember more details than when faced with low- or high-stress situations (Bahrick, Parker, Fivush, & Levitt, 1998). While this finding suggests that the amount of stress may impact upon how the report is relayed, this does not give information as to how true and false reports of stressful events would differ. Similarly, studies attempting to obtain measures of children's reports of stressful events that considered visits to a doctor (Pezdek et al., 2004) or to the emergency room (Peterson, McDermott Sales, Rees, & Fivush, 2007) did not compare children's true and false reports of such events. To date, no published studies have compared children's true and false reports of victimization, which is one of the most forensically significant contexts.

While children have been found to be capable of lying, a number of studies have revealed that when adults attempt to differentiate between children's true and deceptive statements, they are highly inaccurate and rarely perform above chance levels (Crossman & Lewis, 2006; Edelstein, Luten, Ekman, & Goodman, 2006; Leach, Talwar, Lee, Bala, & Lindsay, 2004; Orcutt et al., 2001; Strömwall, Bengtsson, Leander, & Granhag, 2004; Strömwall et al., 2007; Talwar & Lee, 2002; Talwar et al., 2006; Tye et al., 1999; Vrij, Akehurst, Brown, & Mann, 2006). This is consistent with the finding of a recent meta-analysis that lay adults' detection of other adults' deception is near chance levels (C. F. Bond & DePaulo, 2006).

Methods of Veracity Detection

Because untrained adults have difficulty consistently and accurately detecting deception (C. F. Bond & DePaulo, 2006), more objective and systematic methods of lie detection must be explored. In an effort to increase efficiency and accuracy in veracity classification, methods investigating linguistic differences in true and false stories have been developed. Such methods, for example criteria-based content analysis (CBCA) and reality monitoring, are based on the Undeustch hypothesis, which posits that true and false stories are, by nature, different (Undeutsch, 1982, p. 44). By detecting the presence or absence of certain variables, both CBCA and reality monitoring have been shown to have success in determining truths and lies, with CBCA demonstrating classification rates ranging from 65 to 90% (see Vrij, 2005 for review) and reality monitoring obtaining classification rates ranging from 64 to 86% (see Sporer, 2004 for review). Although these rates are significantly above chance levels, these methods have limitations which may reduce their utility. For example, these methods are drastically affected by the length of the statement, which is problematic as

younger children's reports tend to contain fewer details and are generally shorter (Goodman & Reed, 1986; Pipe, Lamb, Orbach, & Esplin, 2004; Vrij, 2005; Vrij, Akehurst, Soukara, & Bull, 2004b). Further, both methods of classification require extensive training, are labour-intensive to complete and are typically subjective.

Advances in technology may assist with credibility assessment through the practical application of more objective methods to detect differences in language between truths and lies. Automated linguistic analysis software programs allow for an investigation into semantic speech patterns. One such program is the Linguistic Inquiry Word Count (LIWC) computer software program, which detects semantic patterns of speech by analysing text and calculating the frequencies of word use in relation to the total word count (Pennebaker, Francis, & Booth, 2001). While this method has the potential to assist in advancing knowledge about the composition of truthful and fabricated statements, the software's utility for classifying the veracity of children's reports has rarely been tested. Evans and colleagues (2012) used LIWC to consider linguistic differences between children's true and false stories about playing a game, and considered how language changed with repetition. Differences were found between true and false stories, with Sensory and Perceptual details occurring more frequently in false than in true stories (see Evans et al., 2012).

Semantic analysis using LIWC has been applied to adult accounts of true and false reports and reveals several linguistic patterns. For instance, adults' false reports tend to contain fewer exclusive terms (e.g., except, without), more relativity terms (motion and spatial terms), fewer first person pronouns and fewer negative emotional words (G. D. Bond & Lee, 2005; Newman, Pennebaker, Berry, & Richards, 2003). As exclusive terms can signify complexity and relativity terms may be easier to discuss, it is possible that these linguistic patterns of false reports are semantically simpler and are employed strategically to decrease cognitive demands. This is consistent with other research which suggests that deception requires more cognitive effort and higher processing abilities than telling the truth (Sporer & Schwandt, 2006; Vrij, 2000). Further, by using fewer self references and more motion terms, the individual may be creating distance between themselves and the lie, and redirecting the listener's attention (G. D. Bond & Lee, 2005; Newman et al., 2003). While this computer-based method increases the speed and objectivity in the linguistic analysis of witness's statements, few studies have applied such technology to children's reports of events.

In this study, we applied the LIWC software to analyse the linguistic patterns in 7–14-year-old children's true and fabricated reports of both stressful and non-stressful events. Reports of stressful events in the current study required children to discuss one true and one fabricated situation in which they had been bullied by another child. Bullying that involves a threat of physical harm or the application of force is a criminal act andmay cause victims to suffer emotionally or physically. As the damaging effects of bullying are increasingly recognized, victims of bullying are being encouraged to come forward and to report incidents. A child who has been bullied may experience stress, embarrassment, shame or feelings of responsibility, which may replicate some of the feelings children experience when reporting about child abuse or neglect (Bottoms & Goodman, 1994; Goodman, Bottoms, Schwartz-Kenney, & Rudy, 1991; Smith, 2008). Thus, bullying is a naturally occurring phenomenon and allows for an ethical investigation into how legally relevant events are reported and fabricated by children.

Research on bullying reveals that this is a widespread problem, with estimates of school-aged children who have been bullied ranging from 15 to 45% (Craig, 1998; Veenstra et al., 2005). However, fewer than half of bullying victims report the incident, potentially out of fear that they will not be believed (Theriot, Dulmus, Sowers, & Johnson, 2005). Also of concern is the potential for bullying claims to be false. While no formal studies have

examined false bullying claims, organizations such as "Bullies to Buddies" have claimed that with anti-bullying campaigns in schools, students are increasingly reporting false incidents in order to get other students in trouble (Kalman, 2005). When incidents of bullying are reported, parents, teachers, school administrators, and increasingly police and youth court judges must determine if children's reports are honest and subsequently establish an appropriate course of action. An improved understanding of differences between true and false reports of such bullying events would assist in accurate classification and response to these reports.

The study also included children's reports of a true and fabricated non-stressful sporting event. This control condition was intended to provide a point of comparison and to evaluate whether stressful and non-stressful events truly differ in how they are delivered. Children's verbal statements concerning both true and fabricated stressful and non-stressful events were analysed for linguistic differences and to determine if any semantic patterns could be discerned.

Based on the existing findings regarding a developmental trajectory in children's reports of events, younger participants were expected to have shorter reports (Goodman et al., 1991). It was also expected that subtle linguistic differences would be revealed between true and false reports; similarly to adult reports (G. D. Bond & Lee, 2005; Newman et al., 2003), false reports were expected to be shorter, have fewer self references, use more relativity terms (such as motion terms and spatial terms) and display linguistic patterns that decreased cognitive demands (such as fewer discrepancy terms, fewer exclusive terms and more tentative terms), when compared with true reports. It was also expected that children would make fewer self references in the stressful bullying stories compared with non-stressful sport stories in an attempt to distance themselves from the potential embarrassment of the story topic. Finally, given the hypothesized differences between true and false reports, we expected that both stressful and non-stressful true and false reports would be able to be statistically discriminated above chance levels.

Method

Participants

A total of 42 participants between the ages of 7 to 14 ($M = 10.52$, $SD = 2.30$; 16 females) were interviewed in two large urban cities in Canada. Participants were divided into four age groups: 7–8-year-olds ($N = 9$), 9–10-year-olds ($N = 13$), 11–12-year-olds ($N = 9$) and 13–14-year-olds ($N = 11$). Participants were recruited through newspaper advertisements and flyers in the community that specifically called for child participants who had been bullied. Parents were instructed to discuss the study with their children prior to participating, but asked not to prepare specific stories with their children in advance. Assent was obtained from all children prior to participating in the study.

Procedure and Materials

Upon arrival, participants were informed that they would be telling two stories that would be true and two that would be fabricated, for a total of four stories. Two stories were about a sports event, with one true account and one fabricated. Similarly, two stories were about a bullying event, with one being true and one being fabricated. To begin, participants told their true sports story to a research assistant and notes were made concerning the main points of the story. The research assistant then facilitated the creation of a false story concerning a sporting event using general prompts (e.g., What sport would you like to talk about? What happened? Who was there?). Further, children were asked how similar this story was to something that had happened to them before, and then asked to identify which

parts were true and which parts were false. Over half of the children (56.1%) indicated that the story was not similar to something that had happened to them before, while 24.4% indicated it was a little similar. Further investigation into the true aspects of the false story revealed that children indicated concrete details, such as names or places, which were similar to true events, but that the events themselves were different.

The child then discussed a true bullying event with the researcher. The research assistant ensured that the child was comfortable discussing the material and informed participants that they could stop at any point if necessary. No children stated that they wanted to stop or that they were uncomfortable with the procedure. A fabricated bullying event was created in the same fashion as the false sport story, and the child was asked to make up a story about a type of bullying that they had never experienced (e.g., physical assault, verbal threat of assault, extortion, exclusion, gossip). Like the false sports story, the research assistant facilitated the creation of a false story concerning a bullying event using general prompts (e.g., What happened? Who was there?). Again, children were asked how similar the fabricated story was to something they had actually experienced; 51.2% of participants indicated the false story was not similar to something they had experienced and 29.3% indicated it was a little similar. Children again frequently used true concrete details, such as names or places, but the general content of the story was fabricated.

Because planning can moderate deception abilities (Sporer & Schwandt, 2006, 2007), each story was created in the same manner and told only once to the research assistant to control for practice effects prior to the final report made to the interviewer. Such methods more closely replicate a police investigation where a child will usually have discussed their report with an adult, like a teacher or parent, before making a statement to the police.

Following this practice phase, participants were individually taken to the testing area with a second research assistant (the interviewer). Participants were seated in front of a camera on a tripod. The interviewer sat behind the camera and discussed the videotaping procedure with the child. The interviewer had notes from the "practice" session about which sporting events and which bullying events to ask the child about. After discussing the procedure with the child, the camera was turned on and the child was asked a general prompt to begin (e.g., "Tell me about the time you played baseball"). When the child was finished recounting details about the event, the research assistant asked an openended prompt, "Is there anything else you remember?" or "How did you feel?". The same process was repeated for each of the four stories. The entire procedure was completed in a single half-hour session.

In order to increase the participant's comfort, the story order was held constant with true and false non-stressful sport stories told first, followed by true and false stressful bullying stories. The selected order was due to the emotional and stressful content of the true bullying stories and allowed for an opportunity for rapport and trust to build between the participant and the research assistants. For their participation, children received ten dollars.

Linguistic Pattern Detection

Each story was transcribed word-forword, and prepared according to the requirements identified by the LIWC 2001 Manual (Pennebaker et al., 2001). Only the participants' statements, not interviewer questions, were included in the subsequent analysis. The four stories told by each child were processed by the LIWC linguistic software system, which classifies each word within a specific linguistic category. When using all 72 categories, LIWC accurately categorizes 80% of all words (Pennebaker et al., 2001). For the purposes and content of the study, 29 categories were used for the LIWC analysis. The majority of the words in each statement were categorized, with no significant differences between story types. In line with the procedure used by Newman et al. (2003), categories with low

frequencies (< 0.2% of the time), or categories left up to the discretion of the transcriber (e.g. non-fluencies and fillers, such as "umm" or "uh") were excluded. In total, 19 variables were included in the subsequent analysis (see Table 1). With the exception of word count, the frequency of each category was obtained as a percentage of total words for each story type.

Results

Preliminary analysis revealed no significant differences between genders for any of the LIWC categories. Thus, the data of both genders were collapsed for the subsequent analyses.

To evaluate whether children's true and fabricated reports differ in various linguistic categories, a series of 2 (Story Content: stressful vs. non-stressful) × 2 (Veracity: true vs. false) × 4 (Age Groups: 7–8-year-olds, 9–10-year-olds, 11–12-yearolds and 13–14-year-olds) mixed analyses of variance (ANOVAs) were conducted for the 19 LIWC categories. See Table 1 for a complete list of means and standard deviations by Story Content and Veracity. We begin by examining the length of report followed by the linguistic differences.

Length of Report

The length of each story (indexed by the word count variable), revealed a significant main effect of Story Content, $F(1, 38) = 16.37$, $p < .001$, $\eta^2 = .30$, 95% confidence interval (CI .95) = [35.76, 107.36], indicating that more words were spoken in the stressful compared with the non-stressful stories (see Table 2 for means and standard deviations). A significant interaction of Veracity and Age Group was also found, $F(1, 38) = 5.95$, $p = .002$, $\eta^2 = .32$. To investigate this interaction, four pairedsample t-tests were performed between the true and false stories for each age group. Significant results were found for 9- and 10-year-olds only and revealed that significantly more words were in their true stories compared to their false stories, $t(12) = 3.82$, $p < .001$, CI .95 = [81.37, 297.86].

Linguistic Variables

Results revealed several significant differences in terms of the types of words participants used in their stories. First, we report significant differences that were found between the stressful and nonstressful stories (Story Content). Next, we report significant main effects of Veracity as well as significant interactions between Veracity and Story Content. Finally, we report on significant interactions between Veracity and Age. No significant three-way interactions between Story Content, Veracity and Age were found.

Significant Effects of Story Content—For several linguistic variables, significant main effects of Story Content were found, which illustrates that the language used in stressful and non-stressful stories was linguistically different. As predicted, children's non-stressful stories were found to have significantly more positive emotional words and significantly fewer negative emotional words than stressful stories, $F(1, 38) = 32.09$, $p < .001$, $\eta^2 = .49$, CI .95 = [1.17, 2.47], and $F(1, 38) = 19.71$, $p < .001$, $\eta^2 = .34$, CI .95 = [.38, 1.03], respectively. Children were also found to use significantly more words that were six letters long (or more) in their stressful stories compared with non-stressful stories, $F(1, 38) = 9.80$, $p = .003$, $\eta^2 = .21$, CI .95 = [.43, 2.00]. Further, stressful stories contained more sensory and perceptual process words and more tentative terms than non-stressful stories, $F(1, 38) = 109.10$, $p < .001$, $\eta^2 = .74$, CI .95 = [1.61, 2.39], and $F(1, 38) = 10.61$, $p = .002$, $\eta^2 = .22$, CI .95 = [.27, 1.15], respectively.

Significant Effects of Veracity—When considering differences in terms of Veracity, the ANOVA yielded a significant main effect when considering self references $F(1, 38) =$

155.21, $p < .001$, $\eta^2 = .80$, CI .95 = [3.85, 5.34], with more self references being made in false rather than true stories. Further, when considering self references, there was a significant interaction between Veracity and Story Content, $F(1, 38) = 107.14$, $p < .001$, $\eta^2 = .74$. Four followup paired-sample t-tests comparing true, false, stressful, and non-stressful stories revealed significantly more self references in false than in true stressful and nonstressful stories, ($t(41) = 2.58$, $p = .01$, CI .95 = [−1.77, −.21], and $t(41) = 14.23$, $p < .001$, CI .95 = [−9.49, −7.13], respectively). In addition, there were more self references in true stressful than in true non-stressful stories, $t(41) = 13.10$, $p < .001$, CI .95 = [−7.99, −5.86].

A significant interaction between Story Content and Veracity was also found for spatial terms, $F(1, 38) = 5.14$, $p = .03$, $\eta^2 = 12$. Four follow-up paired-sample t-tests between the true stories, the false stories, the non-stressful stories, and stressful stories revealed significantly more spatial terms in true non-stressful stories compared with true stressful stories, $t(41) = 2.32$, $p = .03$, CI .95 = [.09, 1.24].

Finally, the ANOVA yielded a significant interaction between Veracity and Age Groups for motion terms, $F(3, 38) = 3.56$, $p = .02$, $\eta^2 = .22$. Four follow-up pairedsample t-tests between the true and false statements for each age group revealed significant results for 9- and 10-year-olds only and found more motion terms in their false stories compared to their true stories, $t(12) = 2.71$, $p < .02$, CI .95 = [−1.40, −.15].

Discriminant Analysis

As ANOVAs revealed limited linguistic differences between true and false sport and stressful stories, discriminant analyses using the Wilks' lambda method were conducted to determine whether the veracity of children's statements could be predicted using these linguistic trends. Discriminant analyses were completed for both stressful and nonstressful stories separately in order to determine if the emotional content of the stories impacted how accurately the veracity of the stories could be classified.

Two separate analyses were run for each story type, with the veracity of the statement entered as the classifying variable. The first discriminant analysis considered the four variables that had significant main effects or interactions with Veracity as a result of the ANOVAs, which were age groups, word count, self references and motion terms. As the ANOVAs were based on group differences for individual variables, information regarding patterns of differences among the variables may not be evident by only considering the significant variables resulting from the ANOVAs. Thus, in addition to the four variables identified above, a step-wise discriminant analysis was used to identify whether additional linguistic variables could enhance veracity classification accuracy. Thus, all linguistic variables that were initially considered (listed in Table 1) were included in a second discriminant analysis.

Non-stressful Stories—When considering the first four variable discriminant analysis for non-stressful stories, the overall Wilks' lambda was significant, Wilks' $\Lambda = .32$, $\chi^2(5, N=84) = 89.65$, $p < .001$, indicating that true and false non-stressful stories could be successfully differentiated above chance levels (50%). Specifically, these factors accurately predicted true and false stories 92.9% of the time (91.7% cross-validation group classification). A closer analysis of the structure matrix revealed that the self reference variable ($\beta = 1.00$) was a significant predictor above and beyond the common contributions of all variables. This finding suggests that children use more self references in their false reports compared to true non-stressful stories. No other predictors were uniquely significant.

When considering all 29 variables in the step-wise analysis, results did not reveal any additional significant predictors of veracity classification accuracy for the non-stressful stories.

Bullying Stories—The first step-wise discriminant analysis with the four variables for the stressful stories was not significant, Wilks' $\Lambda = .92$, $\chi^2(5, N=84) = 6.95$, $p = .59$, with only 64.2% of the stories being classified correctly (52.4% cross-validation group classification). No variables were found to be significant unique predictors of veracity. However, when entering all 29 linguistic variables as predictors using the step-wise method, the tentative terms variable was identified as a significant predictor, making the overall Wilks' lambda significant, Wilks' $\Lambda = .95$, $\chi^2(1, N = 84) = 4.46$, $p = .04$, with 59.9% of stories (59.9% cross-validation group classification) accurately classified. An analysis of the structure matrix revealed that the use of tentative terms ($\beta = 1.00$) was a significant predictor of veracity classification above and beyond the common contributions of all linguistic variables. Children were found to use more tentative terms in their true stressful accounts than in their false stressful stories.

Discussion

The study examined the classification of children's true and false reports of stressful and non-stressful events. To date, few studies have made comparisons between children's true and false reports and to our knowledge children's fabricated reports of stressful victimization and non-stressful events have not been directly compared. Results of the present study revealed significant linguistic differences between children's true and false statements, as well as between stressful and non-stressful reports of events.

Stressful Bullying vs Non-stressful Sport Stories

The current study provides unique insight into children's reports of both stressful and non-stressful stories as the LIWC analysis reveals evidence for differences in linguistic patterns based on story content. Unsurprisingly, the stressful stories were found to contain more negative emotional words and fewer positive emotional words than the non-stressful stories. Further, more spatial terms were used in true non-stressful stories than in true stressful stories. Considering that the non-stressful stories involved descriptions of sports events, the use of spatial terms may be more related to the topic than to the non-stressful nature of the story. Given the significant differences in the use of emotional words, results provide evidence that the story conditions were significantly different.

The structure of children's statements also appeared to be impacted by the emotional content of the stories. Overall, stressful stories contained more words, suggesting that children were capable of discussing and articulating reports than contained emotional information. This finding is consistent with literature indicating that children's descriptions of emotional incidents are more detailed and potentially longer than non-emotional accounts (Goodman, Batterman-Faunce, Shaaf, & Kenney, 2002; see Pipe et al., 2004 for a review). Further, children's bullying stories contained, on average, more words that were at least six letters long. This suggests that children's stressful reports contained larger and possibly more sophisticated words.

In addition, more sensory and perceptual terms and more tentative terms were found in the stressful stories than in the non-stressful stories. These findings suggest that the emotional content of the stories may impact the language used; describing stressful events may elicit more reliance on one's external experience and may also cause one to be more hesitant while relaying the information. Current findings suggest that additional sensory information was found generally in stressful stories, suggesting that emotional content, not veracity, may

impact the occurrence of such types of linguistic cues. Further, children's reports of stressful events increased their use of tentative terms, which may cause the child to appear uncertain and to not be believed.

When considering the results of the discriminant analysis, the difference in the veracity classification accuracy suggests that children may be better able to fabricate stressful stories to resemble true stressful stories than with non-stressful content as there are more distinct linguistic differences between true and false non-stressful stories. One possible explanation for this finding is that after experiencing the emotional event of being bullied, children ruminate and can replicate their emotions and feelings more comprehensively when fabricating an account. Children who have been bullied have knowledge of details affiliated with such incidents and in order to tell more sophisticated lies, they draw on experience from similar events to make their report more believable. In this way, the stressful bullying stories may be more formulaic and as a result, these children's knowledge of similar events may make it easier for them to fabricate accounts that closely resemble true events. Indeed, familiarity of the event has been shown to improve one's ability to create a more believable false story (Blandon-Gitlin et al., 2005; Pezdek et al., 2004). Future studies including participants who have not been bullied or experienced the stimulus stressful event in any way are needed to determine if the ability to closely replicate true stories is a result of experience.

True vs Fabricated Stories

When considering differences in true versus false reports, the self references variable was significantly different for both stressful and non-stressful stories, but was only a significant predictor for non-stressful stories. Contrary to our hypothesis and adult literature, children referred to themselves more in their false stressful reports than in their true stressful reports. By referring to themselves more often during false stories, children may be decreasing the cognitive demands of the task as it is easier to discuss oneself and one's actions than those occurring around them. Further, children may be using fewer self references in their true stressful accounts in order to distance themselves from the traumatic or stressful event. Overall, findings suggest that children use different strategies when fabricating stories than adults do. Therefore, generalizations from adult literature should be applied with great caution and more studies considering the developmental trends in story fabrication should be considered.

Evidence that true stories would contain more complex semantic terms, such as discrepancy and exclusion terms, was not found. However, the follow-up discriminant analysis provided evidence that children use more tentative terms in their true stressful reports than in their false stressful reports. This has significant implications when considering children's stressful accounts as the use of tentative terms may make the child appear apprehensive and may suggest that they are not telling the truth. However, it appears as though the stressful content of the true stories makes children more hesitant when recounting emotional events and tentativeness actually may indicate truth. How the use of tentative language impacts upon adult detection rates is currently unknown.

The results also revealed that children's true stories were in fact longer than their false stories, although this effect was only evident for 9- and 10-year-olds. Further, there was a significant difference in motion terms where true non-stressful stories contained more than true stressful stories, but this finding was again qualified by age and only found with the 9- and 10-year-olds. This provides some evidence for the hypothesis that false stories would contain more relativity terms, such as motion terms, which may decrease cognitive demands, but was not found to be significant across age groups.

Because 9- and 10-year-olds performed differently from other age groups, the reasoning behind this must be considered. One possibility is that this age group was the most motivated to do well in the present task. The older age groups may not have been as motivated, and therefore did not expand on or put as much effort into constructing their responses. The younger age group may have been more hesitant to share their responses. As motivation has been found to be a moderating factor in deception (Sporer & Schwandt, 2006; 2007), future studies in which motivation is manipulated or controlled for are needed. It must also be considered that the 9- and 10-year-olds are the largest age group ($N = 13$) and this could be driving the results due to issues with power.

Limitations

A unique component of this study is the ecological relevance of using bullying, a naturally occurring stressful event that involves victimization of children, and in many cases, reports about crimes. However, the nature of such studies raises questions about consistency in children's reports; it would be unethical to create stressful situations and ask children to report on them, raising the issue of establishing "ground truth", which is the knowledge that the statement is actually true or false (Vrij & Mann, 2004). As ground truth is not guaranteed when describing naturalistic events, the implications this has for this study must be considered (Vrij, 2000; Vrij & Mann, 2004). The methodology of the study attempted to control ground truth in various ways as parents contacted experimenters (indicating that parents had knowledge of their child's victimization), children were assented individually, and they did not have to participate in the study and were told they would still be compensated. Therefore, there was no motivation to fabricate a story and pretend it was true. Similarly, we cannot be certain that children's "false" stories did not actually occur. As described in the methods, children were asked to identify how similar their false story was to something that has happened to them and to identify the ways in which it was similar. While the majority of children reported that it was not similar or was a little similar to events that had actually happened, the degree of similarity is beyond control of the study. Future studies may attempt to further control the veracity and accuracy of story content by asking parents or teachers to describe the bullying situation and compare it with the child's report, or have the parent present while the child is reporting their story. Such controls may allow for exaggerated truths or inaccuracies to be rectified.

The ability to detect children's deceptive statements about victimization and other stressful events has potential implications for the legal system and the professionals who work in it, such as police officers, lawyers, child protection workers and judges. Children and adolescents are often witnesses in courtroom situations and are required to testify about their own abuse, neglect or victimization and the outcomes of such cases have serious implications. Further, those working with children in non-forensic situations, such as teachers and social workers, require methods for discriminating between true and fabricated reports of events in order to respond appropriately. The current study provides a starting point for further research that may eventually be used to support professionals in detecting the truthfulness of children's statements.

Acknowledgments

References

Bahrick LE, Parker JF, Fivush R, Levitt M. The effects of stress on children's memory for a natural disaster. Journal of Experimental Psychology. 1998; 4:308–331.

Bala N, Ramakrishnan K, Lindsay R, Lee K. Judicial assessment of the credibility of child witnesses. Alberta Law Review. 2005; 42:995–1017. Retrieved from http://heinonline.org/HOL/LandingPage?collection=journals&handle=hein.journals/alblr42&div=43&id=&page=journals&handle=hein.journals/alblr42&div=43&id=&page=.

Blandon-Gitlin I, Pezdek K, Rogers M, Brodie L. Detecting deception in children: an experimental study of the effect of event familiarity on CBCA ratings. Law and Human Behaviour. 2005; 29:187–197.

Bond CF, DePaulo BM. The accuracy of deception judgements. Personality and Social Psychology Review. 2006; 10:213–234.

Bond GD, Lee AT. Language of lies in prison: linguistic classification of prisoners' truthful and deceptive natural language. Applied Cognitive Psychology. 2005; 19:313–329.

Bottoms BL, Goodman GS. Perceptions of children's credibility in sexual assault cases. Journal of Applied Social Psychology. 1994; 24:702–732.

Bruck M, Ceci SJ. The suggestibility of children's memory. Annual Review of Psychology. 1999; 50:419–439.

Craig W. The relationship among bullying, victimization, depression, anxiety, and aggression in elementary school children. Personality and Individual Differences. 1998; 24:123–130.

Crossman AM, Lewis M. Adults' ability to detect children's lying. Behavioural Sciences and the Law. 2006; 24(4):703–715.

Evans AD, Brunet MK, Talwar V, Bala N, Lindsay RCL, Lee K. The effects of repetition on children's true and false reports. Psychiatry, Psychology and Law. 2012; 19(4):517–529.

Edelstein RS, Luten TL, Ekman P, Goodman GS. Detecting lies in children and adults. Law and Human Behaviour. 2006; 30:1–10.

Goodman GS, Batterman-Faunce JM, Shaaf JM, Kenney R. Nearly 4 years after an event: children's eyewitness memory and adults' perceptions of children's accuracy. Child Abuse and Neglect. 2002; 26(8):849–884. [PubMed: 12363335]

Goodman GS, Bottoms BL, Schwartz-Kenney BM, Rudy L. Children's testimony about a stressful event: improving children's reports. Journal of Narrative and Life History. 1991; 1:69–99.

Goodman GS, Reed DS. Age differences in child witness testimony. Law and Human Behavior. 1986; 10:317–332.

Kalman, IM. A new kind of bullying: bearing false witness. Bullies to Buddies: A Psychological Approach to Bullying, Newsletter. 2005. Retrieved 6 October 2009 from http://www.bullies2buddies.com/A-New-Kind-of-Bullying-Bearing-False-Witness

Leach A, Talwar V, Lee K, Bala N, Lindsay RCL. Intuitive lie detection of children's deception by law enforcement officials and university students. Law & Human Behavior. 2004; 28:661–685. [PubMed: 15732652]

Lewis, M. The development of deception. In: Lewis, M.; Saarni, C., editors. Lying and deception in everyday life. New York: Guilford Press; 1993. p. 90-105.

Lewis M, Stanger C, Sullivan MW. Deception in 3 year olds. Developmental Psychology. 1989; 25:439–443.

Lyon T, Malloy L, Quas JA, Talwar V. Coaching, truth induction, and young maltreated children's false allegations and false denials. Child Development. 2008; 79:914–929. [PubMed: 18717898]

Newcombe PA, Bransgrove J. Perceptions of witness credibility: variations across age. Journal of Applied Developmental Psychology. 2007; 28:318–331.

Newman ML, Pennebaker JW, Berry DS, Richards JM. Lying words: predicting deception from linguistic styles. Society for Personality and Social Psychology. 2003; 29:665–675.

Orcutt HK, Goodman GS, Tobey AE, Batterman-Faunce JM, Thomas S. Detecting deception in children's testimony: factfinders' abilities to reach the truth in open court and closed-circuit trials. Law and Human Behavior. 2001; 25:339–372. [PubMed: 11501438]

Pennebaker, JW.; Francis, ME.; Booth, RJ. Linguistic Inquiry and Word Count (LIWC): LIWC 2001 manual. Mahwah, NJ: Erlbaum; 2001.

Peskin J. Ruse and representations: on children's ability to conceal information. Developmental Psychology. 1992; 28:84–89.

Peterson C, McDermott Sales J, Rees M, Fivush R. Parent–child talk and children's memory for stressful events. Applied Cognitive Psychology. 2007; 21:1057–1075.

Pezdek K, Morrow A, Blandon-Gitlin I, Goodman GS, Quas JA, Saywitz KJ, et al. Detecting deception in children: event familiarity affects criterion based content analysis ratings. Journal of Applied Psychology. 2004; 89:119–126. [PubMed: 14769124]

Pipe ME, Lamb ME, Orbach Y, Esplin PW. Recent research on children's testimony about experienced and witnessed events. Developmental Review. 2004; 24:440–468.

Quas JA, Goodman GS, Ghetti S, Redlich AD. Questioning the child witness: what can we conclude from the research thus far? Trauma, Violence, & Abuse. 2000; 1:223–249.

Smith MC. Pre-professional mandated reporters' understanding of young children's eyewitness testimony: implications for training. Children and Youth Services Review. 2008; 30:1355–1365.

Sporer, SL. Reality monitoring and the detection of deception. In: Granhag, PA.; Strömwall, LA., editors. The detection of deception in forensic contexts. Cambridge: Cambridge University Press; 2004. p. 64-102.

Sporer SL, Schwandt B. Para-verbal indicators of deception: a meta-analytic synthesis. Applied Cognitive Psychology. 2006; 20:421–446.

Sporer SL, Schwandt B. Moderators of nonverbal indicators of deception: a metaanalytic synthesis. Psychology, Public Policy and Law. 2007; 13:1–34.

Strömwall LA, Bengtsson L, Leander L, Granhag PA. Assessing children's statements: The impact of a repeated experience on CBCA and RM ratings. Applied Cognitive Psychology. 2004; 18:653–668.

Strömwall LA, Granhag PA, Landström S. Children's prepared and unprepared lies: can adults see through their strategies? Applied Cognitive Psychology. 2007; 21:457–471.

Talwar V, Gordon HM, Lee K. Lying in the elementary school years: verbal deception and its relation to second-order belief understanding. Developmental Psychology. 2007; 43:804–810. [PubMed: 17484589]

Talwar V, Lee K. Development of lying to conceal a transgression: children's control of expressive behaviour during verbal deception. International Journal of Behavioral Development. 2002; 26:436–444.

Talwar V, Lee K. Social and cognitive correlates of children's lying behavior. Child Development. 2008; 79:866–881. [PubMed: 18717895]

Talwar V, Lee K, Bala N, Lindsay RCL. Children's conceptual knowledge of lying and its relation to their actual behaviors: implications for court competence examinations. Law and Human Behavior. 2002; 26:395–415. [PubMed: 12182530]

Talwar V, Lee K, Bala N, Lindsay RCL. Adults' judgements of children's coached reports. Law and Human Behavior. 2006; 30:561–570. [PubMed: 16941237]

Talwar V, Lee K, Bala N, Lindsay RCL. Children's lie-telling to conceal parents' transgressions: legal implications. Law and Human Behavior. 2004; 28:411–435. [PubMed: 15499823]

Theriot MT, Dulmus CN, Sowers KM, Johnson TK. Factors relating to self-identification among bullying victims. Children and Youth Services Review. 2005; 27:979–994.

Tye MC, Amato SL, Honts CR, Devitt MK, Peters D. The willingness of children to lie and the assessment of credibility in an ecologically relevant laboratory setting. Applied Developmental Science. 1999; 3:92–109.

Undeutsch, U. Statement reality analysis. In: Trankell, A., editor. Reconstructing the past: The role of psychologists in criminal trials. Stockholm, Sweden: Norstedt & Sons; 1982. p. 27-56.

Veenstra R, Lindenberg S, Oldehinkel AJ, De Winter AF, Verhulst FC, Ormel J. Bullying and victimization in elementary schools: a comparison of bullies, victims, bully/victims, and uninvolved preadolescents. Developmental Psychology. 2005; 41:672–682. [PubMed: 16060813]

Vrij, A. Detecting lies and deceit: The psychology of lying and the implications for professional practice. Chichester, UK: Wiley; 2000.

Vrij A. Criteria-based content analysis: a qualitative review of the first 37 studies. Psychology, Public Policy and the Law. 2005; 11:3–41.

Vrij A, Akehurst L, Brown L, Mann S. Detecting lies in young children, adolescents and adults. Applied Cognitive Psychology. 2006; 20:1225–1237.

Vrij A, Akehurst L, Soukara S, Bull R. Detecting deceit via analyses of verbal and nonverbal behavior in children and adults. Human Communication Research. 2004a; 30:8–41.

Vrij A, Akehurst L, Soukara S, Bull R. 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 Behavioral Science. 2004b; 36:13–126.

Vrij A, Mann S. Detecting deception: the benefit of looking at a combination of behavioral, auditory and speech content related cues in a systematic manner. Group Decision and Negotiation. 2004; 13:61–79.

Table 1

Frequency of Linguistic Inquiry Word Count categories by story type.

Dimension	Example	True stressful		False stressful		True non-stressful		False non-stressful	
		M	SD	M	SD	M	SD	M	SD
Linguistic dimensions									
Categorized words		45.12	11.56	44.43	9.10	43.28	9.38	42.50	10.38
Word count		315.67	240.48	258.26	181.00	220.33	136.81	212.10	157.86
Six-letter words		9.71	2.36	9.29	2.67	7.65	2.41	9.00	3.09
First-person singular	I, my	8.89	2.42	9.97	2.70	9.18	3.8	8.25	3.38
First-person plural	we, us	0.85	1.02	0.76	0.87	2.82	2.74	2.88	3.04
Total self references	I, we	9.75	2.09	10.74	2.47	2.82	2.74	11.13	3.08
Second person pronouns	you, your	0.62	0.81	0.61	0.82	0.39	0.83	0.25	0.54
Third person pronouns	she, their	6.21	2.02	5.97	2.55	2.19	1.53	2.87	2.71
Psychological processes									
Positive emotions	happy, pretty	1.64	1.16	1.75	1.32	3.44	2.06	3.62	2.11
Positive feelings	joy, love	0.43	0.87	0.26	0.32	0.42	0.52	0.31	0.46
Negative emotions	hate, enemy	2.10	1.17	1.91	1.15	1.07	0.88	1.53	1.3
Sensory/perceptual Processes	see, touch	3.46	1.33	3.56	1.48	1.44	1.28	1.57	1.26
Tentative terms	maybe, guess	2.75	1.68	2.06	1.21	1.80	2.02	1.64	1.41
Cognitive processes	know, ought	5.60	2.89	5.45	2.56	4.58	2.47	5.08	2.51
Certainty	always, never	0.66	0.71	0.76	0.68	0.59	0.79	0.75	0.85
Discrepancy terms	should, would	1.85	1.37	1.79	1.45	1.23	1.36	1.67	1.39
Relativity									
Spatial terms	around, over	2.61	1.49	3.02	1.52	3.28	1.62	2.81	1.70
Motion verbs	walk, move	1.55	1.14	1.95	1.34	2.16	1.74	1.87	1.36
Time	hour, day	6.05	3.06	5.76	2.55	6.31	3.03	6.04	2.88

Table 2

Significant linguistic differences between story type.

Linguistic category	Significant group differences		
	M (SD)		*M (SD)*
	Stressful		Non-stressful
Main effect of story content			
Word count	286.96 (201.09)	>	216.21 (138.67)
Positive emotions	1.70 (1.03)	<	3.53 (1.60)
Negative emotions	2.01 (.95)	>	1.30 (.82)
Six-letter words	9.50 (1.96)	>	8.32 (1.73)
Sensory and perceptual processes	3.51 (1.04)	>	1.50 (1.03)
Tentative words	2.40 (1.18)	>	1.72 (1.44)
Veracity by Age interactions	True		False
Word count (9–10-year-olds only)	742.15 (395.15)	>	552.54 (333.93)
Motion terms (9–10-year-olds only)	1.21 (.83)	<	1.99 (.84)
Veracity by Story Type interactions	True stressful		True non-stressful
Spatial terms	2.61 (1.49)	<	3.28 (1.62)
Motion terms	1.56 (1.14)	<	2.16 (1.74)
Self references	9.75 (2.09)	>	2.82 (2.74)
Self references	False stressful 10.74 (2.47)	>	True stressful 9.75 (2.09)
Self references	True non-stressful 2.82 (2.74)	<	False non-stressful 11.13 (3.08)

Note: The arrow indicates the direction of the significant differences between groups.