

## Identifying Liars Through Automatic Decoding of Children's Facial Expressions

Kaila C. Bruer 

*University of Toronto and University of Regina*

Xiao Pan Ding

*National University of Singapore*

Kang Lee 

*University of Toronto*

Sarah Zanette

*University of Toronto*

Thomas D. Lyon

*University of Southern California—Gould School of Law*

This study explored whether children's ( $N = 158$ ; 4- to 9 years old) nonverbal facial expressions can be used to identify when children are being deceptive. Using a computer vision program to automatically decode children's facial expressions according to the Facial Action Coding System, this study employed machine learning to determine whether facial expressions can be used to discriminate between children who concealed breaking a toy (liars) and those who did not break a toy (nonliars). Results found that, regardless of age or history of maltreatment, children's facial expressions could accurately (73%) be distinguished between liars and nonliars. Two emotions, surprise and fear, were more strongly expressed by liars than nonliars. These findings provide evidence to support the use of automatically coded facial expressions to detect children's deception.

Understanding why people deceive and when we are being deceived has held our fascination for over a century (see Darwin, 1877). This fascination exists, in part, because deception is a pervasive and universal human behavior that signifies an important part of typical human development (e.g., Lewis, Stanger, & Sullivan, 1989; Talwar & Crossman, 2011; Talwar & Lee, 2008; for a review see Lee, 2013). Children as young as 2 years of age are capable of verbally deceiving others (Evans & Lee, 2013; Lewis et al., 1989; Talwar & Lee, 2002)—though the quality and sophistication of these early lies are quite limited (Evans & Lee, 2013; Fu, Evans, Xu, & Lee, 2012). Children's ability to tell convincing lies improves rapidly with age due to developments in executive control and theory of mind (Evans, Xu, & Lee, 2011; Talwar & Lee, 2002; see Lee, 2013). By the age of six, children can convincingly deceive others to protect someone's feelings (e.g., Saarni, 1989; Talwar & Lee, 2008; Talwar, Murphy, & Lee, 2007), hide their own bad behavior (e.g., breaking a rule; Evans & Lee, 2013; Evans

et al., 2011; Fu et al., 2012), hide their joint transgressions (Stolzenberg, McWilliams, & Lyon, 2017), or hide the negative behavior of another (e.g., a parent or stranger; Talwar, Lee, Bala, & Lindsay, 2004; Tye, Amato, Honts, Devitt, & Peters, 1999). Whether they be lies of commission (i.e., false reports) or lies of omission (i.e., concealing information), this research has shown that children tell very convincing lies—making it difficult to determine when they are lying or telling the truth. Since children are capable of convincingly deceiving others, many people—including parents, teachers, and police officers—are interested in identifying when children are being deceptive. However, existing research has failed to provide a reliable means to identify when children are being deceitful (Gongola, Scurich, & Quas, 2017). In the present study, we addressed this significant gap using a cutting-edge computer vision technology to detect children's deception.

### *Detecting Childhood Deception*

Deception, including verbal lie-telling, involves intentionally instilling a false-belief in another

This study was supported in part by the National Institute of Child Health and Development (Grant HD087685), the Social Sciences and Humanities Research Council of Canada, and the Canadian Foundation for Innovation.

Correspondence concerning this article should be addressed to Kaila C. Bruer, Department of Psychology, University of Regina, 3737 Wascanad Pkwy, Regina, SK S4S0A2. Electronic mail may be sent to [kaila.bruer@uregina.ca](mailto:kaila.bruer@uregina.ca).

© 2019 Society for Research in Child Development  
All rights reserved. 0009-3920/2020/9104-0036  
DOI: 10.1111/cdev.13336

person by either failing to volunteer information, feigning ignorance, or creating a false statement (Talwar, Lee, Bala, & Lindsay, 2002). To successfully deceive another, one must make the appropriate verbal statements as well as convey nonverbal information (i.e., facial expressions, body language) that matches the social context of the lie being told (e.g., smiling while pretending to like a disappointing gift; Zquette, Gao, Brunet, Bartlett, & Lee, 2016). By 6 years of age, most children have developed the skills necessary to convincingly deceive another by adapting their verbal and nonverbal behaviors (e.g., Evans & Lee, 2013; Lewis et al., 1989; Talwar & Lee, 2002). As a result, even individuals highly trained in detecting deception, such as police and customs officers, struggle to identify when children (or adults) are being deceptive—including lies of omission (Crossman & Lewis, 2006; Leach, Talwar, Lee, Bala, & Lindsay, 2004; Talwar, Lee, Bala, & Lindsay, 2006). One reason why adults struggle to detect deception is due to the reliance on inappropriate or inaccurate cues (Akehurst, Köhnken, Vrij, & Bull, 1996) and errors in human-judgment—including an inflated confidence in one's own lie-detection abilities (Leach et al., 2004). Another reason is that children may learn to tell lies differently because of their past experiences. For instance, victims of sexual abuse, physical abuse, or neglect may have a greater history of hiding negative experiences from others (e.g., hiding maltreatment) compared to typical children (Hershkowitz, Lamb, & Katz, 2014). Having more opportunities to conceal negative experiences may result in maltreated children developing more sophisticated lie-telling abilities. Furthermore, given that maltreated children tend to anticipate adults to be unsupportive and punitive (e.g., Shields, Ryan, & Cicchetti, 2001), they may also be more likely to lie to adults about other nonmaltreatment-related events as well (Talwar & Lee, 2011; Williams, McWilliams, & Lyon, 2019). Taken together, such prior experience may impact how maltreated children tell lies and, in turn, our ability to detect their lies.

Our inability to reliably detect children's deception can have a variety of negative consequences. These consequences may be relatively benign, such as not knowing when a child is lying about having brushed their teeth or completed a household chore, or they could be more significant, such as failing to identify when a child is hiding a negative experience such as bullying or maltreatment. An inability to detect children's deception within the legal system has particularly severe implications. Each year, many children are interviewed by police

and later asked to testify in court as witnesses to crimes (e.g., domestic violence), victims of maltreatment (Bala, Lee, Lindsay, & Talwar, 2000), or during a civil (e.g., custody) disputes (Goodman, Quas, Bulkley, & Shapiro, 1999). In any of these situations, it is conceivable that a child may be asked to lie by an adult (Conte, Wolf, & Smither, 1989; Pipe & Goodman, 1991). Without the ability to identify children who are being deceptive, we risk exposing children to continued stress, criminal activity, or even abuse. Therefore, there is a critical need to identify those children who may be deceiving us.

Being able to identify children's deception is important for understanding children's overall social development—including moral development in particular. Using social domain theory, when assessing whether an act is a moral violation or not, children will, with increasing age, differentially consider the context of the event (see Smetana, Jambon, & Ball, 2014), which can be used to understand age-related changes when children will (or will not) lie. For example, as children get older (10- to 11 years old) there is a documented shift away from truthfully reporting most transgressions and toward only reporting more serious transgressions (e.g., Loke, Heyman, Forgie, & McCarthy, 2011). When a child witnesses (or even participates) in a transgression, but is then asked not to tell (i.e., keep it a secret), a child may experience competing moral concerns associated with both keeping the secret as well as for disclosing the secret. Utilizing laboratory-based paradigms to explore children's behavior when they are asked to lie at the request of an adult presents a unique opportunity to study deception and its relation to the development of morality.

Many different methods of deception detection have been explored with children. Verbal cues (e.g., slower speech, more pauses, and more speech disturbances) have been documented as cues to children's deception (Vrij, Akehurst, Soukara, & Bull, 2004). However, relying on children's verbal descriptions when identifying deception can be problematic, because children provide less verbal information than adults do, which limits the potential for analysis (e.g., Vrij et al., 2004). Research has also considered children's nonverbal behaviors as a way to detect deception, either alone or in conjunction with verbal behavior (e.g., Talwar & Lee, 2002; Vrij, Edward, Roberts, & Bull, 2000). For children, nonverbal behaviors are thought to be especially informative because they are more difficult to control compared to the content of speech.

Facial expressions, in particular, have been argued as a possible nonverbal cue for deception

given that young children (< 9 years old) have underdeveloped cognitive capacities (i.e., inhibitory control, theory of mind; Evans et al., 2011; Talwar & Lee, 2008). Furthermore, young children are not able to manipulate their facial musculature as well as adults (e.g., Ekman, Roper, & Hager, 1980; Feldman, Jenkins, & Popoola, 1979), which limits how much control they can exert over their facial expressions. Therefore, using facial expressions as a means of detecting deception, if reliable, would be highly beneficial—particularly because coding facial expressions requires no physical contact with the potential deceiver and is noninvasive compared to existing methods (e.g., polygraphs and other physiological measures). Identifying the expression of emotions from facial expressions may be particularly informative, since nonverbal behaviors vary depending on current emotions (Ekman, 1992). The emotion(s) expressed during lie-telling may provide context to how children navigate between two potentially competing moral outcomes. That is, examining children's facial expressions when telling a lie can help us understand the extent to which children can control their facial expressions to tell a convincing lie.

#### *Automated Decoding of Facial Expressions*

A prominent method used by researchers to measure facial expressions is the Facial Action Coding System (FACS), which measures human facial muscle movements (Ekman, Friesen, & Hager, 2002). FACS is traditionally coded using a labor-intensive hand-scoring method (Ekman & Friesen, 1978). Using this human-coding system, researchers have found that young children who lied about peeking at a toy displayed more frequent or bigger smiles compared to nonliars (Lewis et al., 1989; Talwar & Lee, 2002).

Although this hand-coding FACS classification system has been demonstrated to be reliable and is widely used (Ekman & Rosenberg, 1997), it is very expensive and inefficient (Ekman & Friesen, 1978); Hand-coding requires many hours of intensive training, can take up to 3 hr to manually code 1 min of behavior, and is susceptible to human-related errors stemming from attention-lapses and fatigue (Ekman & Friesen, 1978; Littlewort, Whitehill, et al., 2011; Zanette et al., 2016). Thankfully, with the advancement of computer vision technology and machine learning, the FACS coding method has now been automated (e.g., Bartlett et al., 2006; iMotions, 2018; Lewinski, den Uyl, & Butler, 2014). Using video files, live video, or

individual images, automated FACS technology conducts a frame-by-frame analysis to quickly and reliably code facial muscle movements (i.e., action units; Ekman & Friesen, 1978) and simultaneously generate emotion expression scores (e.g., joy, sadness, anger, disgust; iMotions, 2018).

There are many advantages to using this automated FACS technology. First, it is capable of retroactively investigating facial expressions using only a simple video recording—allowing for a non-invasive method to examine facial expressions in a variety of settings (e.g., Bartlett, Littlewort, Frank, & Lee, 2014; Zanette et al., 2016). Second, it is faster and more cost-effective than the traditional hand-coding method. Third, this automated technology is less susceptible to typical human-coder issues, making it a significantly more reliable option as well (Littlewort, Whitehill, et al., 2011). Fourth and most importantly, automated FACS methods can analyze facial expressions frame-by-frame, allowing for greater coding precision than human-coding and for discovering even fleeting facial expression changes.

Automated FACS technology has been widely used with adult populations and has been shown to code facial and emotion expressions reliably and accurately (e.g., Bartlett et al., 2006; Littlewort, Bartlett, & Lee, 2009). Automated FACS coding technologies have also been used with children (e.g., Littlewort, Bartlett, Salamanca, & Reilly, 2011); however, only one study has applied this method to children's deception. Using automated FACS software, Zanette et al. (2016) decoded children's facial expressions and associated emotions and found that children's facial expressions could distinguish between children telling prosocial lies (i.e., "white lies") from those telling antisocial lies. Due to the fact that all children in their study lied, Zanette et al. (2016) could not address the question of whether the automated facial expression could distinguish between liars and nonliars. As such, it is entirely unknown as to whether this technology can be used to distinguish between children who are being deceptive from those being honest. We examined this possibility in the present study.

#### *Present Study*

The goal of the current study was to determine whether the facial expressions of children (4- to 9 years old) telling lies of omission (not disclosing a transgression) differ from the facial expressions of children not telling a lie (control group). We used two novel methods to examine whether children's

facial expressions could be used to separate liars from nonliars: First, we used automated FACS technology to analyze the facial expressions of children previously recorded in two prior studies (see Procedure for full description). Next, we used machine learning to systematically test whether this automated facial expression data could be used to distinguish liars from nonliars (i.e., detect deception). Previous research using hand-coding methods provide evidence that children's facial expressions do in fact differ during deception compared to when being honest (Talwar & Lee, 2002). Based on the existing, albeit limited research, we hypothesized that automated facial expression data would accurately classify liars from nonliars above chance level.

We were also interested in examining which specific emotion expressions differentiate liars from nonliars. Previous research has found that smiling (Talwar & Lee, 2002) and joyful expressions (Zanette et al., 2016) are related to children's lie-telling; However, these studies focused on children's lies about breaking a rule (i.e., temptation resistance paradigm). The present study is the first to examine children's facial expressions while lying at the request of an adult confederate, as well as the first to examine facial expressions during lies of omission. As such, the exact emotion expressions that would play important roles in this process was largely an exploratory research question.

We also investigated whether the relation between specific emotions and lie-telling varied depending on the children's age or gender. We chose to examine 4- to 9-year-old children because existing studies have shown that by this point in their development most children have refined their lie-telling abilities to such a degree that most adults cannot distinguish their lies from general observation (Gongola et al., 2017) or based on facial expressions (see Lee, 2013). This creates an ideal situation to assess whether automatically coded facial expressions can be used to identify liars where humans cannot. In addition, our sample was evenly split between girls and boys, allowing us to examine potential gender-related differences. Previous research (Talwar & Lee, 2002) has not found age- or gender-related differences in specific emotions that distinguish liars from nonliars. However, we sought to reexamine this research question using automated FACS software to see if automatically coded facial expressions (coded in more detail than previous research) would reveal age- or gender-related differences that previous hand-coding research was not able to detect (Talwar & Lee, 2002). As such, our predictions were largely exploratory in nature

and we made no specific hypotheses regarding whether we would find evidence of age- or gender-related differences in how children's facial expressions are uniquely expressed in liars.

Lastly, in addition to examining potential age- and gender- related differences, we also investigated whether the relations between specific emotions and lie-telling varied depending on whether children have a known history of experiencing maltreatment. In the current study, half of the children had experienced a history of maltreatment (abuse or neglect; see full description as follows). As previously discussed, due to their previous experiences and likely expectations of adults being unsupportive, we hypothesized that maltreated children would lie differently (i.e., display more levels of negative expressions, such as fear, contempt, or disgust) than nonmaltreated children.

## Method

### *Participants*

A total of 158 children aged 4- to 9 years old who were recruited during prior studies (McWilliams, Stenzenberg, Williams, & Lyon, 2019; Stolzenberg et al., 2017) were included in this study. An additional 30 participants were excluded due to issues related to coding of facial expressions (discussed below). Despite the loss of these 30 participants, the sample size of 158 children is estimated to produce stable results for the present study (Janitza & Hornung, 2018; see Data Analysis description) and is nearly three times the size of the previously published work using similar methods (Zanette et al., 2016).

Sixty-one percent ( $N = 93$ ) of children had a history of maltreatment (hereby referred to as the maltreated group), which in this study refers to children who were substantiated as neglected or abused and were removed from the custody of their parents or guardians ( $M_{\text{age}} = 6.66$ ,  $SD = 1.61$ ; 49% females; 69% Latino/a, 16% African American, 9% other, 6% biracial). For these children, consent for participation was obtained from the Presiding judge of Juvenile Court. Thirty-nine percent ( $N = 65$ ) of children were recruited from local schools or aftercare programs and had no known history of maltreatment (hereby be referred to as the nonmaltreated group;  $M_{\text{age}} = 6.57$ ,  $SD = 1.47$ ; 54% females; 82% Latino/a, 7% African American, 10% biracial). Consent for participation was obtained from the children's parent. Recruitment was conducted in a way that ensured the



demographics of the nonmaltreated group (e.g., ethnicity, socioeconomic status) were similar to those in the maltreated group. Specifically, children in the nonmaltreated sample were recruited from schools serving predominantly ethnic minority families in neighborhoods comparable to those from which most maltreated children were removed. As a proxy for socioeconomic-status of nonmaltreated participants, an average of 92% of children qualified for free or reduced lunch programs within their schools and, as such, are likely from low-income households. For the maltreated group, previous investigations with a similar sample suggests that they are also low income, determined by 86% of mothers previously receiving financial aid and food stamps (Lyon & Saywitz, 1999). In all cases, children also assented to participation and were fully debriefed at the end of the procedure.

Children participated in a video-recorded interview as part of a prior research study (McWilliams et al., 2019; Stolzenberg, et al., 2017). In the McWilliams et al. (2019) study, children played with several toys in a research laboratory with an adult confederate. While interacting with the confederate, two of the toys were manipulated so that they appeared to break when played with. The confederate then asked children to lie about this event to conceal the breakage. During a subsequent interview, children had the opportunity to either lie about the toy breaking or disclose the negative event. For the present study, we sought to analyze the facial expressions of all children who engaged in a lie of omission by choosing not to disclose the breaking of toys. These children formed our experimental group (i.e., liars). The protocol used in this prior study required that all children experience the toy breaking. Therefore, we also used videos of children who participated in a second similar study, but for whom no toys broke (Stolzenberg et al., 2017). These children formed our control group (i.e., nonliars) to compare against the lie-tellers. All children in the present study were asked the same open-ended question about their time with the confederate. Children's facial expressions during the interview were recorded using a hidden camera.

#### *Procedure*

Children worked with a female interviewer individually in a quiet room. After completing additional tasks unrelated to the current study, the female interviewer excused herself from the room and left the child alone with boxes of toys. Next, a second person, a female confederate, entered the

room and expressed interest in playing with the toys. What happened next depended on which condition children were in. For children in the experimental group (67%,  $n = 106$ ), two toys were manipulated so that they appeared to break when the child played with them. The confederate then asked the child not to tell the first woman that they had broken the toys, warning that "we might get in trouble if she finds out." The confederate then left the room and the interviewer returned shortly thereafter. Following a brief rapport-building phase, the interviewer asked children a free recall question (i.e., *tell me everything that happened when the lady came in while I was gone*), followed by related follow-up questions not discussed in the present article. The sample selected for analysis included all children in the experimental condition who failed to reveal that the toys had broken. These children are subsequently referred to as the *liars*. For the children in the control condition (33%), no toys broke while the child and confederate played with them. Thus, when asked the open-ended invitation question, children in the control group had no negative event to conceal. These children are therefore referred to as *nonliars*.

Unbeknownst to children, their facial expressions during the interview were recorded using a hidden video camera (parents or guardians were informed of this during consent). The videos were then analyzed using the automatic computer vision technology to obtain facial expression scores that were then used for analysis.

#### *Facial Expression Data*

We used an automated FACS software, FACET (originally the Computer Expression Recognition Toolbox or CERT; iMotions, 2018), that was originally developed by Bartlett et al. (2006). FACET is a fully automatic and real-time software tool for coding facial expressions (Bartlett et al., 2006; see Littlewort, Whitehill, et al., 2011, for an overview). Using computer vision technology, FACET generates a frame-by-frame analysis of the extent to which each facial muscle (i.e., action unit) is activated. Then, using machine learning, FACET translates these muscle movements into a measure of emotion expression. This translation process is done using multinomial logistic regressions that are trained using the Cohn-Kanade data set, a large high-quality database containing thousands of images of each type of emotional expression. The output for each emotion is then passed through a softmax competition (for a complete description, see Kanade, Cohn,

& Tian, 2000; Littlewort, Whitehill, et al., 2011). The final output is a standardized score (logarithmic base 10 scale) reflecting the likelihood that a human coder would identify a set of muscle movements as an emotion, with higher scores indicating that an expression in a frame is more likely to be categorized as an emotion. For example, an anger score of 1 indicates that the observed expression is 10 times more likely to be categorized by an expert human coder as *anger* than *not anger*, while a score of  $-1$  indicates the expression is 10 times more likely to be to be categorized by an expert human coder as *not anger* than *anger*. In the present study we used all 10 emotion expression scores available from FACET: joy, sadness, anger, disgust, surprise, fear, confusion, frustration, contempt, and neutral (see Figure 1).

#### Video Preparation and Data Cleaning

We analyzed children's facial expressions during two key periods of the interview: (a) while hearing the free recall question being asked by the interviewer; and (b) immediately after the free recall question was asked, but before the child provided a

verbal response. We examined these two parts of the interview separately because children may be engaging in different cognitive tasks during each time period. This, in turn, could differently impact facial expressions across these two sections of the interview. That is, the first period may reflect the time when children were first made aware that they would be asked to talk about what happened during the visitation. During this stage, children may be more interested in learning what they will be asked about and feeling out the situation than monitoring their facial expressions. Facial expressions during this initial period might be especially informative of lie-telling because children's inhibitions might not be engaged yet. Once they have heard the entire question asked, children's thoughts may then shift to deciding what to say to the interviewer and how they should frame their response. For children in the experimental group, this could reflect the time when they decided to tell a lie by omission instead of revealing that the toys had broken. The children's facial expressions when they were forming a response, in turn, may capture hidden emotion that a child is actively working to suppress

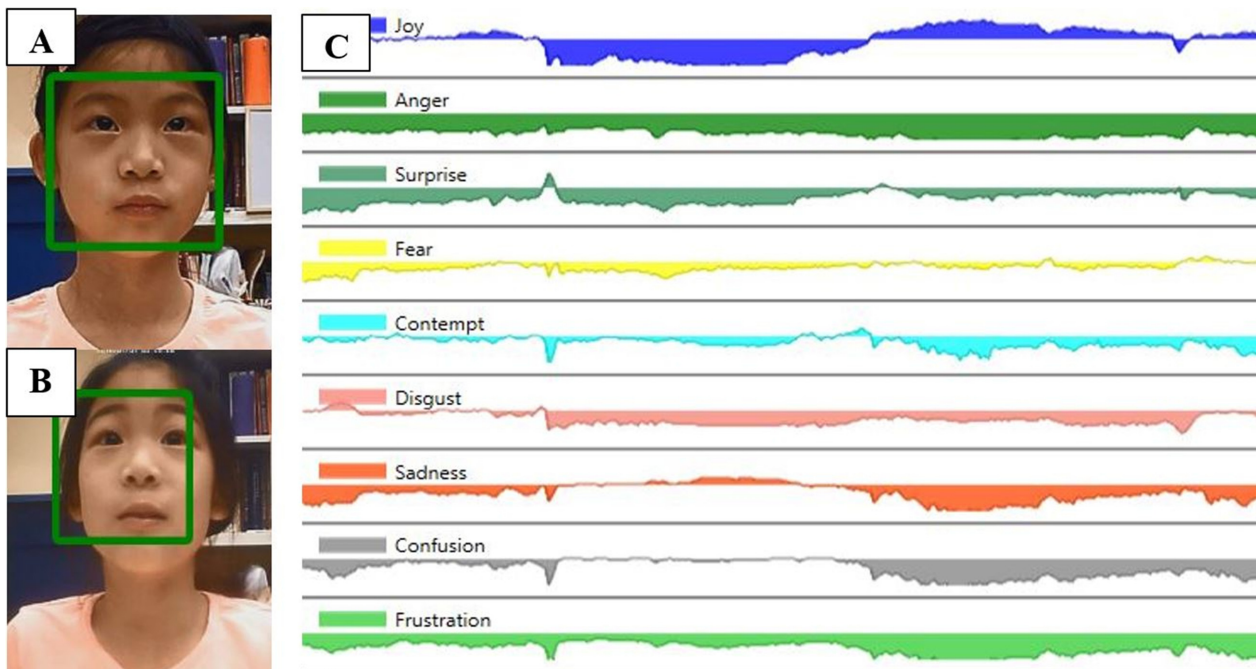


Figure 1. Sample image of FACET and associated emotion expression score output. FACET first locates the face (indicated by the square overlaying the faces in images A and B) and then calculates emotion scores for each frame of the video. Image A shows a child displaying a neutral expression. Image B shows a child displaying the expression of surprise. Image C shows graphs of the emotion scores in FACET. The *x*-axis depicts each frame of video in temporal sequence and the *y*-axis represents the emotion score (positive and negative) with a zero baseline. For example, positive scores on the joy scale indicate that a human coder is more likely to identify that expression as *joy* than *not joy*. In contrast, negative joy scores indicate that a human coder is more likely to identify that expression as *not joy* than *joy*. The children depicted in the figure and their parents have given informed consent to use their likeness in this article.

during lie-telling. Many of the questions asked after this initial free recall question (i.e., cued-initiation questions) varied for each child depending on their earlier responses and because of the changes inherent in natural conversation. Thus, the first free recall question provided the best opportunity to analyze children's facial expressions in a systematic manner with the least amount of variability in procedure across participants.

As performed by Zanette et al. (2016), we ended the segment of video right before the child began to talk to avoid facial expressions being contaminated by the muscle movements involved in speech. Since videos were recorded as part of a natural conversation, the duration of each video segment varied in length and the number of frames (range = 1–484 frames per participant,  $M = 122.11$ ,  $SD = 81.01$ ; recorded at 29 frames per second: range = 0.03–16.69 s;  $M = 4.21$ ,  $SD = 2.79$ ).

A human coder carefully reviewed all video files to removing any facial expression data from video frames containing false-face readings (i.e., FACET incorrectly registering another object in the room as a face), as this would impact the accuracy of the data. Additionally, FACET is unable to measure facial expressions if the child is looking away from the camera or if an object is blocking some or all their face (e.g., the child's hand). These limitations resulted in the loss of facial expression data for some or all video frames recorded for each participant. To manage consistency issues stemming from using participants with missing data, we excluded interviews ( $n = 30$ ) that had more than 50% of frames during the periods of interest register as missing or unreliable (i.e., false-face readings). We then summarized the emotion expression scores for each participant using five metrics: mean, standard deviation, minimum, maximum, and range. This was done separately for each key period of the interview (i.e., while hearing the question and right before responding).

#### *Data Analysis*

Given the highly complex and multidimensional nature of the data, we used a machine learning method referred to as random forest (*randomForest* package, R version 3.4.3, R Core Team, 2017; Liaw & Wiener, 2002; Breiman, 2001) to determine whether the facial expressions could identify liars from nonliars. A random forest model can be conceptualized as a cluster of individual prediction *trees* that form a larger decision *forest*. Individual trees in the model are classification trees that predict the target response—in our case, deception.

Each tree classifies a participant as either a liar or nonliar. The forest classifies individuals by examining the majority vote across the individual tree predictions (500 trees per forest).

Random forest was selected because it is an efficient and robust approach to classification and prediction (Dittman, Khoshgoftaar, Wald, & Napolitano, 2011) and is stable with sample sizes similar to those used in the present study (see Janitza & Hornung, 2018). This is largely due to multiple levels of randomness built into random forest modeling. The first type of randomness is seen during the building of individual trees; random forests uses a “bagging” (i.e., bootstrap aggregation) approach, where the model selects a training set from a random subsection of the data (sampled with replacement). This information is then used to grow (i.e., train and define) each tree. With each subsequent tree grown, the data selected for training will vary. The second level of randomness is that random forest builds trees using information from a subsection of features (i.e., prediction variables) each time. Unlike other models that use all features to build a tree, random forest randomly selects a subset of features to grow each tree. The specific features selected each time varies—however, the number of features is always the same and is determined using the square root of the number of features available. Overall, we chose to use the random forest technique as it creates a more accurate model by averaging several decision trees—this averaging reduces variance in the model and helps to avoid problems stemming from overfitting (Breiman, 2001).

#### *Model Development*

The automated FACS data for the 10 emotions were each summarized into five different metrics: mean, standard deviation, minimum, maximum, and range. To develop a prediction model, we first explored whether the predictive utility of children's emotion expression scores depended on how the data were summarized. We used random forest (*randomForest* package, R version 3.4.3) to conduct a preliminary, but comprehensive analyses to examine which metrics (or combination of metrics), if any, could effectively separate liars from nonliars. We ran a series of prediction models that accounted for (a) the relatively small sample size ( $n = 158$ ) and (b) the unbalanced number of liars ( $n = 106$ ) and nonliars ( $n = 52$ ) in the sample. We evaluated the effectiveness of the different prediction models using three measures: accuracy, sensitivity (*dc*), and response bias (*c*). Accuracy reflects the average

proportion of correct classifications by the model. Sensitivity, or  $d'$ , represents how well the prediction models correctly classify liars relative to misclassifying nonliars as liars, with higher values indicating better discrimination ability (Macmillan & Creelman, 1991). A response bias, or  $c$ , is a measure of bias within the models, with higher values indicating the model is more likely to classify a participant as a liar than a nonliar. The preliminary analyses revealed that using *mean* emotion scores produced the most reliable results (i.e., best able to correctly classify liars from nonliars) compared to the other four metrics used (standard deviation, minimum, maximum, and range). Accordingly, we chose to examine only the results derived from using children's *mean* emotion scores as follows. For a complete discussion of the model development and the preliminary analyses, including the complete results, see Supporting Information (Section S1.1 and S1.2 and appendices).

## Results

Our results are divided into three sections, each answering one of three research questions: (a) Can children's facial expressions, coded using automated FACS technology, accurately distinguish liars from nonliars? (b) If so, which emotion expressions are uniquely expressed in liars and nonliars? (c) How does a child's age, gender, or history of maltreatment impact unique emotion expression during deception? As stated above, all analyses of facial expressions reflect the use of mean scores.

### *Can Children's Facial Expressions Separate Liars From Nonliars?*

Our primary goal was to determine whether automated FACS data can be used to distinguish child

liars from nonliars. As noted above, children's facial expressions were analyzed separately during two response periods: (a) children's facial expressions while hearing the free recall question being asked by the interviewer; and (b) children's facial expressions immediately after the free recall question was asked, but before the child provided a verbal response (i.e., forming a response to the question).

### *Hearing the Question*

Using all 10 emotions, children's facial expressions when hearing the question asked could distinguish liars from nonliars better than chance, with 73% accuracy (see Table 1). However, though we corrected for our sample having more liars than nonliars during model development, this model was still significantly biased toward classifying children as liars over nonliars (response bias score,  $c$ ,  $p < .05$ ). Next, we examined whether the 10 emotions identified liars differently depending on children's age, gender, and history of maltreatment. Using the accuracy of the prediction model as the dependent variable, we ran a 2 (honesty group: liars, nonliars)  $\times$  2 (gender: girl, boy)  $\times$  2 (maltreatment: maltreated, nonmaltreated)  $\times$  2 (model accuracy: accurate, inaccurate) hierarchical log-linear analysis. We also ran a correlation analysis. These results revealed that our prediction model accurately classified liars from nonliars regardless of the child's history of maltreatment,  $\chi^2(1) = 0.001$ ,  $p = .981$ , gender,  $\chi^2(1) = 0.03$ ,  $p = .875$ , or age,  $r(156) = -.04$ ,  $p = .604$ . However, this model was more biased to correctly classify participants as liars (92% accurate classification,  $n = 106$ ) than nonliars (35%,  $n = 52$ ),  $\chi^2(1) = 54.22$ ,  $p < .001$ . These results suggest that the model predicted lie-telling with consistent accuracy, regardless of age, gender, or history of maltreatment.

Table 1  
*Top Performing Models in Predicting Deception*

Part of interview	Model	Proportion accuracy		$d'$		$c$	
		$M$	$p$ Value	$M$	$p$ Value	$M$	$p$ Value
Hearing question	10-feature (Mean)	0.73	< .001	1.00	< .001	0.92	< .001
Forming response	10-feature (Mean)	0.68	< .001	0.45	.904	1.12	< .001

*Note.* For each model, the metric score reflects the average score across 1,000 model generations. The associated  $p$ -value is determined as 1 – number of replications above critical value/number of replications. For example, a  $p$ -value of < .001 indicates that all 1,000 replications were above the critical value, or in the top 2.5% of the associated random distribution. Sensitivity, or  $d'$ , represents a model's discrimination or ability to correctly classify liars relative to misclassifying nonliars as liars, with higher values indicating better discrimination. A response bias, or  $c$ , is a measure of bias within a model, with higher values indicating the model is more likely to classify a participant as a liar than a nonliar.



### Forming a Response

Next, we repeated the same steps to examine children's facial expressions between hearing the question and responding (i.e., forming a response). Table 1 highlights the top performing model during this portion of the interview. Complete model results can be found in S1.2 of the Supporting Information. The top performing model, using the mean score of 10 emotions, classified liars from nonliars with 68% accuracy. However, the model was not able to correctly classify deception better than chance ( $d^2$  scores; see Table 1). Therefore, children's facial expressions while forming a response to the open-ended question were not useful in identifying deceivers. Accordingly, we chose not to further examine facial expressions during this time period.

### Which Emotions Distinguish Liars From Nonliars?

Children's facial expressions when hearing the question, coded using automated FACS technology, accurately distinguished liars from nonliars. But *which* specific emotions are important in this classification process? To explore this, we investigated which emotions were most effective in differentiating liars from nonliars. To do so, we used an adjusted random forest analysis (i.e., *Boruta* package for R; Kursa & Rudnicki, 2010) to examine how well our model predicted deception at above chance levels (at  $p < .05$ ) both with and without each emotion included. For additional details on this analyses, see S.3 in Supporting Information. Two emotions emerged as important in predicting deception: surprise ( $M_{\text{importance}} = 14.20$ ) and fear ( $M_{\text{importance}} = 8.51$ ). See Figure 2 for an approximated visual representation of which mean emotion scores were identified as important predictors of deception.

### Does Gender, Age, or History of Maltreatment Impact Surprise and Fear Expression During Lie-Telling?

Surprise and fear emerged as important emotions in identifying liars. We then explored *how* these two emotions differed across participants by exploring the *mean* emotion expression scores when children heard the question. For both surprise and fear scores, we ran general linear model (GLM) analyses with experimental group (liars vs. nonliars), maltreatment (history of maltreatment, no history of maltreatment), gender (male vs. female), and age (as a continuous variable) as the independent

variables. For each analysis, the false discovery rate correction was applied (Benjamini & Hochberg, 1995).

Results revealed that, as expected, the expressed levels of surprise and fear were significantly different for liars and nonliars (see Figure 3) when hearing the question. For surprise, liars displayed significantly higher evidence of surprise ( $M = -0.36$ ,  $SD = 1.21$ ) than nonliars ( $M = -1.55$ ,  $SD = 1.39$ ),  $F(1, 147) = 28.35$ ,  $p < .001$ ,  $\eta_p^2 = .162$ . By transforming these mean log odd (base 10) scores,<sup>1</sup> we find evidence that there is a 3% chance of nonliars displaying surprise but a 30% chance in liars. For fear, liars displayed significantly higher evidence of fear ( $M = -0.95$ ,  $SD = 0.81$ ) than nonliars ( $M = -1.39$ ,  $SD = 0.85$ ),  $F(1, 147) = 11.17$ ,  $p = .008$ ,  $\eta_p^2 = .071$ . That is, there is a 4% chance of nonliars displaying fear and a 10% of liars displaying fear. Additionally, children with a history of maltreatment displayed higher mean scores of fear ( $M = -0.96$ ,  $SD = 0.89$ ; 10% chance of showing fear) than children with no history of maltreatment ( $M = -1.28$ ,  $SD = 0.74$ ; 5% chance of showing fear;  $F(1, 147) = 6.87$ ,  $p = .039$ ,  $\eta_p^2 = .045$ ). No other significant differences or interactions in mean surprise or fear scores were found (all  $ps > .05$ ). Thus, liars showed greater evidence for surprised and fearful facial expressions compared to nonliars, regardless of their gender, age, or whether they have a history of experiencing maltreatment.

## Discussion

The present study examined whether automatically decoded facial expressions could be used to distinguish children who lied from nonliars. We obtained three main findings: First, children's facial expressions when hearing the question, coded using automated FACS technology, accurately distinguished liars from nonliars. Second, the key emotions that were more strongly expressed in children who lied were surprise and fear. Third, these two emotions were similarly expressed by liars regardless of age, gender, or history of maltreatment. We discuss these findings in the following sequence.

Children's automatically decoded facial expressions while hearing an open-ended, invitation question about their time with the confederate, as well as during the time leading up to their response to the question, were explored as an indicator of

<sup>1</sup>Transformation:  $((10^{\text{original mean score}})/(1 + 10^{\text{original mean score}})) \times 100$ .

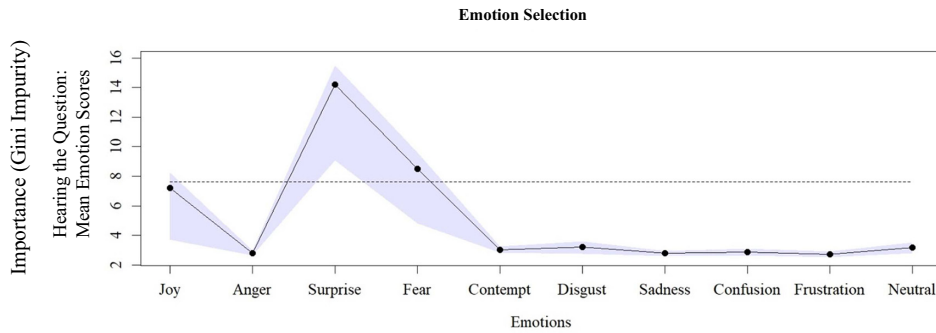


Figure 2. Simplified importance graph. Each point represents the average importance of each emotion as determined by *Baruta* package in R. The shaded area around each point represents the maximum and minimum importance values for each emotion calculated. The dashed line represents the average maximum shadow importance value ( $M = 7.69$ ) that was compared against all other importance values from each emotion.

Note. *Baruta* used  $p < .05$  to determine significance, not these minimum and maximum importance values—however, we include them in the graph for demonstration purposes. For both Surprise and Fear, the average importance was above the maximum shadow importance value, highlighting their importance in predicting deception.

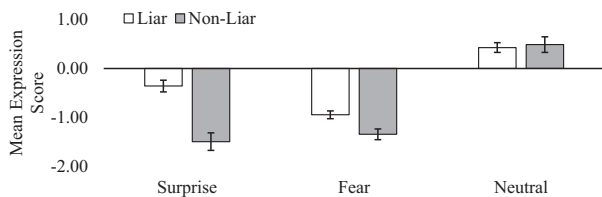


Figure 3. Mean surprise, fear, and neutral emotion expression scores for liars and nonliars when hearing the question.

Note. Neutral is included as a comparison group. Error bars reflect standard error values.

deception. As expected, facial expression scores when hearing the question were accurately and reliably able to separate children who lied about a toy breaking from children who did not lie (i.e., controls), suggesting that children exhibited unique facial expressions when being deceptive. This was true regardless of what metric was used to represent facial expression data (e.g., mean vs. range of scores), the children's gender, or a history of maltreatment. We did, however, find evidence that the metrics used to summarize the automatic FACS data may impact how the data can identify liars across age. When using mean expression scores, the model predicted well regardless of age. However, when using both mean and range scores, the model performed better (i.e., was a better indicator of deception) among younger children. Therefore, how automatic FACS data is summarized should be carefully considered if predicting lie-telling across a wide age range. Overall, our findings provide early, but important evidence that automatic facial expression software can be used to study and detect children's deception.

Surprisingly, children's facial expressions were only useful in identifying liars when first hearing the question—not when framing a response to the question. This suggests that liars displayed different emotions than nonliars did during the time when they were first made aware that the conversation was going to be about the topic they were instructed to be deceptive about. However, it seems that these children were able to quickly control their facial expressions so that no such differences could be detected during the time when they were framing their response to the question. This contrasts with existing research that suggests that children have less control over their facial expressions than adults (e.g., Ekman et al., 1980; Feldman et al., 1979) and instead aligns more with literature suggesting that children as young as four can exert some control over their emotion expressions (Carlson & Wang, 2007). After the topic of conversation was introduced, perhaps children's regulation of emotion expression limited our ability to use facial expressions to detect deception in the second part of the interview. Therefore, the time when the initial question is asked may be the most opportune time to measure children's facial responses, as children may be less guarded when they are not yet aware they will be asked about an event they intend to be deceptive about.

We also predicted that liars would express specific emotions differently than nonliars. We found evidence to support this hypothesis. Unlike existing research that has found positive expressions to be unique to lie-tellers (e.g., smiling; Talwar & Lee, 2002), we found that liars expressed negative emotions differently from nonliars. When children

heard the question, those who lied about the transgression displayed greater mean evidence of surprise and fear than children in the control group.

### *Surprise and Fear During Deception*

Surprise may be differently expressed when children engage in deception. It is also possible that increased expression of surprise in liars may reflect the experimental context of this study. For instance, increased surprise in liars may reflect general arousal when they were asked to think about the transgression (as nonliars did not experience the transgression). Surprise may have been more prominent in liars because this was the first instance that they were asked about the encounter with the confederate toy-breaker. However, other elements of the experiment may have reduced children's surprise about the line of questioning, such as the fact that the transgression had just occurred immediately before questioning, the confederate warned the child that they might get into trouble which a child may have inferred an upcoming questioning, and that the interviewer was someone that they previously met at the start of the study. Additional research is required to speak to whether surprise is consistently expressed by liars across several different contexts.

The higher expression of fear when children lied, however, may be due to the circumstances surrounding the transgression and the confederate's request for the child to be deceptive. In a review of child abuse cases, Hershkowitz, Lanes, and Lamb (2007) reported that half of children described feeling afraid or ashamed of reporting that abuse to their parent when they expected a negative reaction. In the present study, children were primed to anticipate a negative reaction if they told about the broken toys (i.e., getting into trouble). Therefore, the increased levels of fear observed with children who lied about a broken toy, we expect, may have resulted from their belief that they may get into trouble for revealing that the toys broke. Whether fear would continue to be an important factor in identifying deception in situations when children do not anticipate negative responses remains an empirical question for future research.

Moreover, children's unique expression of fear during deception may provide insight into how school-aged children navigate competing moral concerns. Children (even as young as 3 years old; Smetana, 2006) appear to be able to consider the context when assessing whether a moral violation was committed (see Loke et al., 2011; Smetana

et al., 2014). Fear expressions may indicate that children who lied (at the request of an adult) may have been trying to balance competing moral concepts. Liars in our sample were those who experienced the toy breaking and self-selected into the "liars" group by concealing the truth from the interviewer (albeit at the request of the confederate): perhaps children decided to lie due to some perceived social "contract" or expectations of confidentiality with the confederate. However, the display of a fearful expression also suggests that children may perceive their lie as a moral violation, since choosing to lie also breaks a separate social "contract" of honesty during the conversation with the interviewer. This conflict between two social obligations may then manifest itself in the expression of fear. For all these reasons, without replication and experimental consideration of alternative explanations, there is a need to be cautious with interpreting evidence of fear (or even surprise) as a clear indication of deception.

We were also interested in exploring whether these emotions—surprise and fear—expressed during deception varied depending on children's age, gender, or history of maltreatment. As expected, we found no evidence that surprise or fear expression during lie-telling vary by age or gender. This finding aligns with evidence from studies that have explored children's emotions using both hand-coding (Talwar et al., 2007) as well as automatic computer imaging software (Zanette et al., 2016). However, contrary to prediction, we found no evidence that children with a history of maltreatment would lie differently than those without such history. Though maltreated children expressed higher fear scores than nonmaltreated children overall, a history of experiencing maltreatment did not impact children's facial expressions when telling lies. This indicates that the facial expressions of surprise and fear may be useful for identifying deception in children regardless of their age, gender, or history of maltreatment. Replication, including with different experimental designs, will be important next steps.

This research also highlights the possibility of using automated FACS technology to identify potential lie-tellers in applied settings. While we are unable to make definite conclusions as to whether a child is lying at this stage of research, the results of the current study provide evidence of a potential "tell" that occurs among a subset of children, which may be used in future research to refine the use of FACS technology in more applied settings. For example, in legal settings, forensic interviewers are

often interested in identifying victims of maltreatment who may be concealing such experiences. FACS technology may be a means to identify warning signs during an investigation that a child is at risk of concealing a wrongdoing, thereby justifying continued supervision and support, and helping investigators to determine if a subsequent disclosure was due to initial reluctance from the child (Blasbalg, Hershkowitz, Lamb, Karni-Visel, & Ahern, 2019; Saywitz, Wells, Larson, & Hobbs, 2019). However, additional research is needed to refine the use of FACS technology before it can be implemented in real-world settings.

In addition to contributions to our understanding of lying in childhood, this article demonstrates the value of two underutilized research methodologies in developmental science: automated FACS technologies and machine learning. Automated FACS technology is rarely used in the literature but offers significant advantages from both a pragmatic and theoretical perspective. For researchers who are primarily interested in studying children's facial expressions, automated FACS technology offers many practical advantages over traditional hand-coding methods. Not only is automated FACS cost-effective, reliable, and accurate (e.g., Littlewort, Whitehill, et al., 2011), it also allows us to measure intricate differences in facial expression that are beyond the realistic capabilities of human perception. The differences observed in emotion expressions for children who lied was observed at a level that is likely out of reach for human coders (i.e., differences in negative scores of expressions). For instance, nonliars had an average of 3% chance of showing surprise (i.e., almost no evidence of surprise occurring in this group), which was significantly lower than liars who had an average of 30% chance of showing surprise (i.e., some evidence of a surprised facial expression occurring). This means that while the average chance of showing surprise was quite low for both groups, there were still measurable differences in the likelihood that children in each group were showing surprise—differences which would likely go undetected if human coders were used. Children's facial expressions were largely neutral throughout most of the analyzed components of the video (see Figure 3) and, because of this, human coders likely would have only coded neutral (or lack of) emotion expression, which would have completely overlooked the subtle differences among liars and nonliars that FACS technology was able to detect.

Automated FACS programs allow for the collection of facial expression data in a minimally

invasive way—requiring only a video recording of a child's face. This allows for easy integration of this technology into existing research protocols for those developmental researchers who may be peripherally interested in facial expression data. Inclusion of such data may be a method to glean more theoretical insight into the social and cognitive factors that impact behavior. In the present study, this technology, provided insight into what children are experiencing when they engage in deceptive behavior (e.g., surprise and fear emotions), while prior work has used automated FACS to demonstrate that these expressions may differ depending on the type of lie (Zanette et al., 2016). Likewise, clinical (e.g., emotion regulation; assessment or diagnosis), educational (e.g., math anxiety), or sociocognitive researchers may find inclusion of such technology helpful in understanding children's emotional experiences in different settings. For example, researchers have used FACS technology to examine children's facial expressions during problem solving (Littlewort, Whitehill, et al., 2011). Similarly, social domain researchers could use facial expression information to study how children assess the rules governing interpersonal verbal and nonverbal communication in different moral contexts, including lie-telling. Having an objective way to explore what children may be feeling during a task is particularly important when working with young children who may not yet possess the metacognitive skills to reflect upon and report their own emotional experiences (e.g., Flavell, 2000).

Researchers who are interested in using automated FACS technology with videos collected in applied settings should be aware of the potential loss of data due to issues with the quality of footage taken (as experienced in the present study). An important consideration will be to intentionally capture footage that is conducive to this technology, including well-lit environments and subjects that directly face the camera (hidden or otherwise) with nothing blocking their face. Overall, when combined with existing research (e.g., Zanette et al., 2016) our study demonstrates the value that using such technologies can have in studying child development—including children's lying.

A second novel contribution is that this article demonstrates machine learning as a viable approach to study children. Some researchers have used machine learning to identify clinical populations of children (e.g., Duda, Haber, Daniels, & Wall, 2017) but it is rarely used to predict behaviors in children (but see Hagenbuchner, Cliff, Trost, Tuc, & Peoples, 2015). Due to challenges inherent in



working with children (e.g., limited language abilities), experimental paradigms continue to adopt technologies that introduce complex data as a way to obtain objective measures of behavior (e.g., neuroimaging, eye-tracking, biometrics, motion trackers, and reaction times). With this increase in complex data, the need to conduct multidimensional data analyses becomes more prominent. Many developmental researchers currently use GLMs or generalized linear modeling to analyze data. Such data analytic approaches, by their nature, only look at the linear relationships between the predictors and predicted variables (the interactions among the predictors are also linear). These approaches fail to uncover the nonlinear relations between predictor variables and the predicted variable. Machine learning methods often overcome this limitation by identifying more complex and nonlinear relations. As demonstrated in this research, machine learning offers those capabilities to developmental researchers and may be an option to yield a deeper understanding of data.

#### *Limitations and Future Research*

There are some limitations that should be considered when developing future research. First, although our final model performed above chance levels, there is still a great deal of room for improvement. As indicated by our high response bias, our model was less able to accurately predict nonliars (controls). Although we adjusted for unbalanced group size during machine learning training, our persisting bias problem likely stemmed from the lower sample size in the control condition. That is, low frequency of controls may explain why our final model was not well able to predict for this group. Due to our unbalanced sample sizes, the success of the model (73%) does not vastly outperform a simple prediction that all children are liars (106 liars/158 total = 67%). Future researchers should strive to use balanced data to test the replicability of this prediction model. Perhaps with more balanced sample sizes, the predictive utility of facial expression could improve beyond our achieved accuracy of 73%.

Second, we only explored children's facial expressions during a small portion of an interview—when children were asked a free recall question about an event in which a transgression may have transpired. We chose to focus on this element of the interview because such open-ended questions are frequently recommended and used during investigative interviews with children (e.g., Lamb et al.,

2003). However, use of more direct questions about transgressions are documented to impact honesty through increased reports of transgressions (Gordon, Lyon, & Lee, 2014; Thompson, Clarke-Stewart, & Lepore, 1997). For example, Gordon et al. (2014) found that older children were more likely to keep their parent's secret during open-ended recall questions than direct questions (e.g., "What happened to [the toy]?" or "Did your mom/dad ask you not to tell me that she/he broke the toy?"). Perhaps children's facial expressions when hearing more direct questions would produce more discriminating evidence to identify deception. This is worth exploring in future research.

Third, though informative, there are aspects about the experimental design that may not translate well to more applied settings, such as in a classroom, forensic interviews, or courtrooms. For instance, our study involved only one interview session in which a single person interviewed the child. In many applied contexts, the child may be asked different questions, may be questioned multiple times, or be questioned in the presence of multiple people—some of whom may be perceived by child as holding more authority (e.g., teachers, police officers, lawyers). Moreover, these results are limited to children's responses to a relatively mundane transgression (i.e., breaking toys) in which a child was asked to lie by a stranger. However, pressure placed on children to lie by keeping secrets likely varies according to who is making the request (i.e., stranger vs. family or caregiver) and the nature of the transgression. Talwar et al. (2004), for example, examined children's willingness to conceal that their parent broke a toy and found that most (80%) children (3- to 11-year-olds) did not conceal a parent's transgression. On the other hand, Tye et al. (1999) examined children's willingness to lie for a parent in a more serious circumstance—a theft. When the child knew that their parent stole something, nearly 55% of the children concealed what their parent did by accusing the researcher of the study and another 20% lied by making no accusation (i.e., keeping a secret). Moreover, children's facial expressions may be differently predictive depending on the nature of lie being told (e.g., deception about cheating). For these reasons, future research should explore this paradigm in a more naturalistic setting as well as explore how varying the transgressor and the nature of the transgression impacts the predictive utility of children's facial expressions to identify deception.

Fourth, as an initial exploration of this method to study children's deception, we chose to examine

children who had nothing to conceal (true negatives) and children who were concealing a transgression (false negatives). We selected these groups because this is a central concern for forensic interviewers, who wish to identify children who are concealing abuse. However, it does limit the clarity when interpreting the findings. For instance, the increased expression of surprise and fear could be because liars experienced a transgression while non-liars did not. Future research should examine situations where all children experience a transgression and then either conceal it (liars/false negatives) or the truth about it (confessors/true positives) to identify if such emotion expression differences continue to exist. An additional concern, is that children who do disclose may be lying. Thus, future research should also compare true positives and false positives.

### Conclusion

We provide evidence that children who lied in our study expressed different facial expressions than children who did not lie. More specifically, surprise and fear expressions when first hearing the topic of conversation could be used to separate children who were deceptive from those who were not, regardless of the child's age, gender, or history of maltreatment. These findings provide early evidence that a noninvasive automated facial expression tool can be used to signal when children are being deceptive. Important next steps in this research will be to replicate to identify whether these findings are context-specific or reflective of general experiences of lie-telling in childhood.

### References

- Akehurst, L., Köhnken, G., Vrij, A., & Bull, R. (1996). Lay persons' and police officers' beliefs regarding deceptive behaviour. *Applied Cognitive Psychology, 10*, 461–471. [https://doi.org/10.1002/\(SICI\)1099-0720\(199612\)10:6<461:AID-ACP413>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1099-0720(199612)10:6<461:AID-ACP413>3.0.CO;2-2)
- Bala, N., Lee, K., Lindsay, R., & Talwar, V. (2000). A legal and psychological approach to the assessment of the competence of child witnesses. *Osgoode Hall Law Journal, 38*, 409–451.
- Bartlett, M. S., Littlewort, G. C., Frank, M. G., Lainscsek, C., Fasel, I., & Movellan, J. R. (2006). Automatic recognition of facial actions in spontaneous expressions. *Journal of Multimedia, 1*, 22–35. Retrieved from [http://mplab.ucsd.edu/~marni/pubs/Bartlett\\_JMM06.pdf](http://mplab.ucsd.edu/~marni/pubs/Bartlett_JMM06.pdf)
- Bartlett, M. S., Littlewort, G. C., Frank, M. G., & Lee, K. (2014). Automatic decoding of facial movements reveals deceptive pain expressions. *Current Biology, 24*, 738–743. <https://doi.org/10.1016/j.cub.2014.02.009>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*, 289–300. Retrieved from <http://www.jstor.org/stable/2346101>
- Blasbalg, U., Hershkowitz, I., Lamb, M. E., Karni-Visel, Y., & Ahern, E. C. (2019). Is interviewer support associated with the reduced reluctance and enhanced informativeness of alleged child abuse victims? *Law and Human Behavior, 43*, 156–165. <https://doi.org/10.1037/lhb0000303>
- Breiman, L. (2001). Random forests. *Machine learning, 45*, 5–32.
- Carlson, S. M., & Wang, T. S. (2007). Inhibitory control and emotion regulation in preschool children. *Cognitive Development, 22*, 489–510. <https://doi.org/10.1016/j.cogdev.2007.08.002>
- Conte, J. R., Wolf, S., & Smither, T. (1989). What sexual offenders tell us about prevention strategies. *Child Abuse and Neglect, 13*, 293–301. [https://doi.org/10.1016/0145-2134\(89\)90016-1](https://doi.org/10.1016/0145-2134(89)90016-1)
- Crossman, A. M., & Lewis, M. (2006). Adults' ability to detect children's lying. *Behavioral Sciences & the Law, 24*, 703–715. <https://doi.org/10.1002/bsl.731>
- Darwin, C. (1877). A biographical sketch of an infant. *Mind, 2*, 285–294.
- Dittman, D., Khoshgoftaar, T. M., Wald, R., & Napolitano, A. (2011, November). *Random forest: A reliable tool for patient response prediction*. In Bioinformatics and Biomedicine Workshops (BIBMW), 2011 IEEE International Conference on (pp. 289–296). IEEE. <https://doi.org/10.1109/bibmw.2011.6112389>
- Duda, M., Haber, N., Daniels, J., & Wall, D. P. (2017). Crowdsourced validation of a machine-learning classification system for autism and ADHD. *Translational Psychiatry, 7*, e1133. <https://doi.org/10.1038/tp.2017.86>
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion, 6*, 169–200. <https://doi.org/10.1080/02699939208411068>
- Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial action coding system*. Salt Lake City, UT: Network Information Research Corporation.
- Ekman, P., Roper, G., & Hager, J. C. (1980). Deliberate facial movement. *Child Development, 51*, 886–891. <https://doi.org/10.2307/1129478>
- Ekman, P., & Rosenberg, E. L. (Eds.). (1997). *What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system (FACS)*. New York, NY: Oxford University Press.
- Evans, A. D., & Lee, K. (2013). Emergence of lying in very young children. *Developmental Psychology, 49*, 1958–1963. <https://doi.org/10.1037/a0031409>
- Evans, A. D., Xu, F., & Lee, K. (2011). When all signs point to you: Lies told in the face of evidence.

- Developmental Psychology*, 47, 39–49. <https://doi.org/10.1037/a0020787>
- Feldman, R. S., Jenkins, L., & Popoola, O. (1979). Detection of deception in adults and children via facial expressions. *Child Development*, 350–355. <https://doi.org/10.2307/1129409>
- Flavell, J. H. (2000). Development of children's knowledge about the mental world. *International journal of behavioral development*, 24, 15–23. <https://doi.org/10.1080/016502500383421>
- Fu, G., Evans, A. D., Xu, F., & Lee, K. (2012). Young children can tell strategic lies after committing a transgression. *Journal of Experimental Child Psychology*, 113, 147–158. <https://doi.org/10.1016/j.jecp.2012.04.003>
- Gongola, J., Scurich, N., & Quas, J. A. (2017). Detecting deception in children: A meta-analysis. *Law and Human Behavior*, 41, 44–54. <https://doi.org/10.1037/lhb0000211>
- Goodman, G. S., Quas, J. A., Bulkley, J., & Shapiro, C. (1999). Innovations for child witnesses: A national survey. *Psychology, Public Policy, and Law*, 5, 255–281. <https://doi.org/10.1037/1076-8971.5.2.255>
- Gordon, H. M., Lyon, T. D., & Lee, K. (2014). Social and cognitive factors associated with children's secret-keeping for a parent. *Child Development*, 85, 2374–2388. <https://doi.org/10.1111/cdev.12301>
- Hagenbuchner, M., Cliff, D. P., Trost, S. G., Van Tuc, N., & Peoples, G. E. (2015). Prediction of activity type in preschool children using machine learning techniques. *Journal of Science and Medicine in Sport*, 18, 426–431. <https://doi.org/10.1016/j.jsams.2014.06.003>
- Hershkowitz, I., Lamb, M. E., & Katz, C. (2014). Allegation rates in forensic child abuse investigations: Comparing the revised and standard NICHD protocols. *Psychology, Public Policy, and Law*, 20, 336–344. <https://doi.org/10.1037/a0037391>
- Hershkowitz, I., Lanes, O., & Lamb, M. E. (2007). Exploring the disclosure of child sexual abuse with alleged victims and their parents. *Child Abuse & Neglect*, 31, 111–123. <https://doi.org/10.1016/j.chiabu.2006.09.004>
- iMotions. (2018). *iMotions biometric research platform 7.1*. Copenhagen, Denmark: iMotions A/S. Retrieved from <https://imotions.com/facial-expressions/>
- Janitza, S., & Hornung, R. (2018). On the overestimation of random forest's out-of-bag error. *PLoS ONE*, 13, e0201904. <https://doi.org/10.1371/journal.pone.0201904>
- Kanade, T., Cohn, J. F., & Tian, Y. L. (2000). Comprehensive database for facial expression analysis. In T. Kanade, J. F. Cohn, and Y. Tian (Eds.), *Proceedings of the fourth IEEE international conference on automatic face and gesture recognition (FG'00)* (pp. 46–53). Grenoble, France: IEEE.
- Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the boruta package. *Journal of Statistical Software*, 36, 1–13. <https://doi.org/10.18637/jss.v036.i11>
- Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., Stewart, H., & Mitchell, S. (2003). Age differences in young children's responses to open-ended invitations in the course of forensic interviews. *Journal of Consulting and Clinical Psychology*, 71, 926. <https://doi.org/10.1037/0022-006X.71.5.926>
- Leach, A. M., Talwar, V., Lee, K., Bala, N., & Lindsay, R. C. L. (2004). "Intuitive" lie detection of children's deception by law enforcement officials and university students. *Law and Human Behavior*, 28, 661–685. <https://doi.org/10.1007/s10979-004-0793-0>
- Lee, K. (2013). Little liars: Development of verbal deception in children. *Child Development Perspectives*, 7, 91–96. <https://doi.org/10.1111/cdep.12023>
- Lewinski, P., den Uyl, T. M., & Butler, C. (2014). Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader. *Journal of Neuroscience, Psychology, and Economics*, 7, 227–236. <https://doi.org/10.1037/npe0000028>
- Lewis, M., Stanger, C., & Sullivan, M. W. (1989). Deception in three-year-olds. *Developmental Psychology*, 25, 439–443. <https://doi.org/10.1037/00121649.25.3.439>
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2, 18–22. Retrieved from [https://www.r-project.org/doc/Rnews/Rnews\\_2002-3.pdf](https://www.r-project.org/doc/Rnews/Rnews_2002-3.pdf)
- Littlewort, G. C., Bartlett, M. S., & Lee, K. (2009). Automatic coding of facial expressions displayed during posed and genuine pain. *Image and Vision Computing*, 27, 1797–1803. <https://doi.org/10.1016/j.imavis.2008.12.010>
- Littlewort, G. C., Bartlett, M. S., Salamanca, L. P., & Reilly, J. (2011, March). *Automated measurement of children's facial expressions during problem solving tasks*. In Automatic face & gesture recognition and workshops (FG 2011), 2011 IEEE International Conference on (pp. 30–35). IEEE.
- Littlewort, G., Whitehill, J., Wu, T., Fasel, I., Frank, M., Movellan, J., & Bartlett, M. (2011). *The computer expression recognition toolbox (CERT)*. 2011 IEEE international conference on automatic face & gesture recognition and workshops (FG 2011) (pp. 298–305). Santa Barbara, CA: IEEE.
- Loke, I. C., Heyman, G. D., Forgie, J., & McCarthy, A. (2011). Children's moral evaluations of reporting the transgressions of peers: Age differences in evaluations of tattling. *Developmental Psychology*, 47, 1757–1762. <https://doi.org/10.1037/a0025357>
- Lyon, T. D., & Saywitz, K. J. (1999). Young maltreated children's competence to take the oath. *Applied Developmental Science*, 3, 16–27. [https://doi.org/10.1207/s1532480xads0301\\_3](https://doi.org/10.1207/s1532480xads0301_3)
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York, NY: Cambridge University Press.
- McWilliams, K., Stozenberg, S. N., Williams, S., & Lyon, T. D. (2019). Increasing maltreated and nonmaltreated children's recall disclosures of a minor transgression: The effects of back-channel utterances, a promise to tell the truth, and an incremental putative confession. *Child*



- Abuse & Neglect*. Retrieved from <https://ssrn.com/abstract=3381981>
- Pipe, M. E., & Goodman, G. S. (1991). Elements of secrecy: Implications for children's testimony. *Behavioral Sciences & the Law*, 9, 33–41. <https://doi.org/10.1002/bsl.2370090105>
- R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Saarni, C. (1989). Children's understanding of the interpersonal consequences of dissemblance of nonverbal emotional-expressive behavior. *Journal of Nonverbal Behavior*, 12, 275–294. <https://doi.org/10.1007/BF00987596>
- Saywitz, K. J., Wells, C. R., Larson, R. P., & Hobbs, S. D. (2019). Effects of interviewer support on children's memory and suggestibility: Systematic review and meta-analyses of experimental research. *Trauma, Violence, & Abuse*, 20, 22–39. <https://doi.org/10.1177/1524838016683457>
- Shields, A., Ryan, R. M., & Cicchetti, D. (2001). Narrative representations of caregivers and emotion dysregulation as predictors of maltreated children's rejection by peers. *Developmental Psychology*, 37, 321. <https://doi.org/10.1037//0012-1649.37.3.321>
- Smetana, J. G. (2006). Social-cognitive domain theory: Consistencies and variations in children's moral and social judgments. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 119–153). Mahwah, NJ: Erlbaum.
- Smetana, J. G., Jambon, M., & Ball, C. L. (2014). The social domain approach to children's moral and social judgments. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 23–45). New York, NY: Psychology Press.
- Stolzenberg, S., McWilliams, K., & Lyon, T. D. (2017). The effects of the hypothetical putative confession and negatively-valenced yes/no questions on maltreated and non-maltreated children's disclosure of a minor transgression. *Child Maltreatment*, 22, 167–173. <https://doi.org/10.1177/1077559516673734>
- Talwar, V., & Crossman, A. M. (2011). Predictors of children's prosocial lie-telling: Motivation, socialization variables, and moral understanding. *Journal of Experimental Child Psychology*, 110, 373–392. <https://doi.org/10.1016/j.jecp.2011.05.003>
- 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. <https://doi.org/10.1080/01650250143000373>
- Talwar, V., & Lee, K. (2008). Social and cognitive correlates of children's lying behavior. *Child Development*, 4, 866–881. <https://doi.org/10.1111/j.1467-8624.2008.01164.x>
- Talwar, V., & Lee, K. (2011). A punitive environment fosters children's dishonesty: A natural experiment. *Child Development*, 82, 1751–1758. <https://doi.org/10.1111/j.1467-8624.2011.01663.x>
- Talwar, V., Lee, K., Bala, N., & Lindsay, R. C. L. (2002). Children's conceptual knowledge of lying and its relation to their actual behavior: Implications for court competence examinations. *Law and Human Behavior*, 26, 395–415. <https://doi.org/10.1023/A:1016379104959>
- Talwar, V., Lee, K., Bala, N., & Lindsay, R. C. L. (2004). Children's lie-telling to conceal a parents' transgression: Legal implications. *Law and Human Behavior*, 28, 411–435. <https://doi.org/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. <https://doi.org/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. <https://doi.org/10.1177/0165025406073530>
- Thompson, W. C., Clarke-Stewart, K. A., & Lepore, S. J. (1997). What did the janitor do? Suggestive interviewing and the accuracy of children's accounts. *Law and Human Behavior*, 21, 405–426. <https://doi.org/10.1023/A:1024859219764>
- Tye, M. C., Amato, S. L., Honts, C. R., Devitt, M. K., & Peters, D. (1999). The willingness of children to lie and the assessment of credibility in an ecologically relevant laboratory setting. *Applied Developmental Science*, 3, 92–109. [https://doi.org/10.1207/s1532480xads0302\\_4](https://doi.org/10.1207/s1532480xads0302_4)
- Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2004). Detecting deceit via analyses of verbal and nonverbal behavior in children and adults. *Human Communication Research*, 30, 8–41. <https://doi.org/10.1111/j.1468-2958.2004.tb00723.x>
- Vrij, A., Edward, K., Roberts, K. P., & Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal Behavior*, 24, 239–263. <https://doi.org/10.1023/A:1006610329284>
- Williams, S., McWilliams, K., & Lyon, T.D. (2019). Children's concealment of a minor transgression: The role of age, maltreatment, and executive functioning. *Journal of Experimental Child Psychology*. Retrieved from <https://ssrn.com/abstract=3417899>
- Zanette, S., Gao, X., Brunet, M., Bartlett, M. S., & Lee, K. (2016). Automated decoding of facial expressions reveals marked differences in children when telling antisocial versus prosocial lies. *Journal of Experimental Child Psychology*, 150, 165–179. <https://doi.org/10.1016/j.jecp.2016.05.007>

### Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:



**Figure S1.** Histogram Outlining the Distribution of Sensitivity ( $d'$ ) Scores Across the 1,000 Replications of the Top Two Performing Models

**Figure S2.** Emotions Importance by Each Metric

**Table S1.** Top Performing Models in Predicting Deception

**Data S1.** Random Forest Model Development

**Data S2.** Comprehensive Analysis Results

**Data S3.** Bartua Random Forest Model Description

**Appendix S1.** Model Results: Hearing Question

**Appendix S2.** Model Results: Forming Response to Question