Bacso, S., Nilsen, E. S., & Silva, J. (2021). How to turn that frown upside down: Children make use of a listener's facial cues to detect and (attempt to) repair miscommunication. *Journal of Experimental Child Psychology*, 207, 105097. https://doi.org/10.1016/j.jecp.2021.105097

How to turn that frown upside down: Children make use of a listener's facial cues to detect and (attempt to) repair miscommunication

As many parents attest, young children often have difficulty providing sufficient information for their intentions to be clear. There are several decades of studies examining children's ability to refer to objects, people, and events (i.e., referential communication) demonstrating children's initial failures in providing adequate descriptions of target objects for a listener (e.g., Deutsch & Pechman, 1982, Lloyd, Mann, & Peers, 1998). This initial ambiguity makes the ability to detect and repair miscommunication essential for the successful exchange of information, and as such, a key aspect of children's communicative development.

Consider the following example, a child wanting to direct their parent's attention to a car in the distance, says, "Look at the car." even when several cars of different models and colours are on the road, thus leaving their parent unable to effectively understand their communicative intentions. In such an instance, the parent would likely provide the child with verbal cues indicating they failed to understand their request (e.g. saying, "Which one?") as well as nonverbal cues indicating the message was misunderstood (e.g. making a confused facial expression). After receiving this feedback, the child might attempt to repair their message so that it can be successfully understood by the listener (e.g., "The green one").

The process of repairing miscommunication involves several steps, which may be supported by feedback from a communicative partner. First, a speaker must recognize that their communication has been ineffective, a process that can be facilitated through feedback from a communicative partner. Second, they must decide to improve upon their initial statement, and finally, they must generate a repair statement that, if effective, contains additional information to correct the miscommunication. In this way, the first and second steps are necessary pre-requisites to a successful repair, but do not guarantee successful execution.

While past research has explored children's ability to detect and repair miscommunication in response to verbal feedback (e.g., Bacso & Nilsen, 2017; Coon, Lipscomb, & Copple, 1982; Deutsch & Pechmann, 1982; Nilsen & Mangal, 2012; Uzundag & Küntay, 2018), no research has examined this ability in response to purely nonverbal (i.e., affective) feedback. Further, only a handful of studies have sought to determine the cognitive skills that support children's repairs, and none have explored associations with the specific steps involved in children's communicative repairs. The present study addressed these gaps: The first aim was to examine children's ability to detect miscommunication after receiving only nonverbal listener feedback (i.e., a sad facial expression) indicating the child's message was misunderstood, and to, subsequently, repair their messages. The second aim was to examine the role of executive functioning (EF) and emotion knowledge (EK), in children's ability to identify and repair miscommunication in response to nonverbal feedback.

The Role of Feedback in Children's Communication

In general, interactive contexts, where a listener is an active participant in the exchange, are key to children's success in producing clear referential statements (Grigoroglou & Papafragou, 2019). Using cues from a listener, children are able to identify that their own message has been ineffective and is in need of repair. In the absence of any immediate feedback, children do not appear to learn from their mistakes in communication. They tend to persist in producing ambiguous messages, and do not appear to recognize that their original message was inadequate. For instance, Robinson and Robinson (1985) found that 5-year-old children's communicative performance did not improve when the experimenter chose a correct object following ambiguous messages and then subsequently explained what was lacking in the child's message. Similarly, Wardlow and Heyman (2016) found that young school-age children provided more ambiguous descriptions on subsequent trials when an experimenter did not provide feedback versus a condition where feedback was provided. In a context where an experimenter provided an alternate toy to that requested, but no other feedback, toddlers tended to abandon their initial attempts (Fagan, 2008). Thus, feedback appears to be essential for children to learn about the communicative needs of a listener.

With respect to the act of repairing messages specifically, when listeners provide verbal feedback indicating they misunderstood the message, children attempt to repair messages, suggesting that they are attuned to such cues from communicative partners (e.g. Anselmi, Tomasello, & Acunzo, 1986; Bacso & Nilsen, 2017; Coon et al., 1982; Nilsen & Mangal, 2012; Uzundag & Küntay, 2018; Wilcox & Webster, 1980). However, the success of children's repairs depends on the type and quality of feedback provided. For instance, Bacso and Nilsen (2017) found that 4- and 5-year-old children were better able to repair messages that were initially misunderstood by the listener following feedback which specifically identified what was lacking in the child's original message (e.g., "there are three boys and I don't know which one you mean") compared to vague feedback (e.g., "I don't know which one you mean"). Other work has supported this finding that more specific verbal feedback benefits children's communication repair compared to vague feedback (Coon et al., 1982; Nilsen & Mangal, 2012; Uzundag & Küntay, 2018). In addition, repairing statements in response to feedback from a listener may be an important way through which children learn to be effective communicators. Matthews and colleagues (2007; 2012) found that providing preschool-age children with specific verbal feedback, and giving them the opportunity to repair their messages, improved their ability to provide effective descriptions of target objects on subsequent trials during a referential communication task. When preschool-age children are provided with minimal feedback (e.g., "huh" or "what"), they often repeat their original messages rather than attempting to repair them (Anselmi et al., 1986; Nilsen & Mangal, 2012). Thus, it is likely that the type of verbal feedback influences all three steps of the repair process: helping children identify that a message is in need of repair, helping children decide to repair messages, and helping children to produce an effective repair message.

Nonverbal feedback, such as affective cues, may also aid children in recognizing when miscommunication has occurred, although this has not been directly examined. Work by Wardlow and Heyman (2016) provides some support for the notion that children use listeners' emotional cues during referential communication. In this study, 5- to 7-year-old children completed a communication task in which some objects were blocked from the listener's view. When children provided an ambiguous message, the listener either provided no feedback (i.e., smiled and chose the correct object) or provided nonverbal feedback (i.e., looked confused and chose the incorrect object). Children who received the nonverbal feedback provided more effective messages across trials compared to those who received no feedback. However, the fact that the nonverbal feedback used in this study included an incorrect object choice is important, as this feature in isolation is a particularly salient cue for prompting children to repair messages. That is, work by Nilsen and Mangal (2012), found that children often repaired their messages when the listener visibly chose the incorrect object. Thus, it is unclear whether children were responding to the confederate's affective cues or incorrect object choice. The extent to which children can make use of affective cues to guide their ability to produce effective messages for a listener remains unclear.

In addition to the visual aspects, emotional cues exist within the tone of communicators' voices. Past work has found that preschool-age children are sensitive to a speaker's emotional prosody and can use this cue to guide their interpretation of referential communication (see Graham, San Juan, & Khu, 2017 for a review). For example, a study by San Juan and colleagues (2017) found that 5-year-olds were able to use a speaker's happy or sad prosody to determine which of two objects the speaker was referring to when one object was disliked by the speaker, and the other was liked by the speaker. While this line of work suggests that young children can use affective cues when interpreting others' statements, it remains unclear whether they use affective information to *generate* effective messages, and if so, what skills support this process.

Socio-cognitive skills associated with communicative repairs

Given the multiple steps involved in repairs (i.e., detecting miscommunication, deciding to repair, and generating a revised statement), it is likely that there are a number of supporting skills, with the focus in the current study on children's executive functioning (EF) and emotional knowledge (EK). Executive functioning refers to a set of cognitive skills which support goal-directed behavior. The components of executive functioning assessed in the present study include working memory, inhibition, and cognitive flexibility. Emotion knowledge refers to children's ability to recognize emotional expressions, and to understand the situations that elicit emotions (Denham, 1998). Previous research suggests that EF and EK are interrelated, but separable constructs (Martins, Osório, Veríssimo, & Martins, 2016). Further, children with stronger EF skills are likely to develop stronger EK skills over time (See Denham et al., 2012). Both skills show associations with children's socio-communicative behaviour (Bassett, Denham, Mincic, & Graling, 2012; Fabes, Eisenberg, Hanish, & Spinrad, 2001; Matthews, Biney, & Abbot-Smith, 2018), though their role within specific aspects of repair is not known.

While there is some support for the notion that EF is related to children's ability to detect miscommunication (albeit in third person tasks; Gillis & Nilsen, 2014; Nilsen & Graham, 2012), past work in this area has tended to focus on the last step in the repair process (i.e., generating a successful repair). For instance, Wardlow and Heyman (2016) found that children's working memory was associated with their ability to improve their descriptions of target objects across trials when children received feedback on their performance. Interestingly, such relations were not found in a condition where feedback was not provided (similar to other studies showing no relation, e.g., Nilsen & Graham, 2009). Further, Bacso and Nilsen (2017) found that children with higher cognitive flexibility were better able to repair their messages in response to verbal feedback provided by a listener which indicated she had been misunderstood. Work by Uzundag and Küntay (2018) had similar results, finding that cognitive flexibility, working memory, and short-term memory were associated with the quality of children's repairs. It has been reasoned that cognitive flexibility may aid children in repairing their messages by allowing them to view the target object's dimensions more flexibly. For instance, children with stronger cognitive flexibility skills may be better able to identify both the target object's colour *and* shape without becoming fixated solely on the colour of the object. However, it is unclear whether EF shows similar associations across all components of the repair process, particularly in a context where the feedback from the listener involves solely affective information.

Recognizing that a repair is necessary, particularly in a context where explicit verbal cues are not provided, requires that a speaker first detects and recognizes the meaning behind a listener's facial expression. Thus, we would expect that children's ability to identify the need to repair messages and decide to repair messages based on affective cues would relate to their ability to recognize and understand the emotions of others (i.e., EK). Certainly, in contexts outside of communication, preschoolers' EK has predicted their response to others' emotional expressions (e.g., Denham 1986; Denham & Couchoud, 1991).

The current study

The first goal of the present study was to assess 4- and 5-year-old children's ability to use affective cues from a listener to guide their evaluation and repair of messages. We chose this age range due to the rapid growth in communicative ambiguity detection shown within this age range (Nilsen & Graham, 2012) and to be consistent with the literature showing the relations between EF and repairs following *verbal* feedback (e.g., Bacso & Nilsen, 2017; Wardlow & Heyman, 2016). During the current study, children were asked to provide instructions to a (fictional) child listener about where to find a prize. Based on the quality of the children's initial messages, the listener either found the prize and looked happy, or failed to find the prize and looked sad. The listener's affective reaction was matched to the success of the trial to be consistent with how natural interactions unfold.

To capture children's performance on the three steps to the repair process, we asked them to provide perceptions of their communicative success through ratings (detecting miscommunication), decide whether they would like to repair or not (repair decision), and assessed their success at generating a revised statement (repair). We expected that children would demonstrate an ability to use the affective cues of the listener to guide their ratings and to guide their communicative behaviour. The second goal was to examine the associations between children's EF and EK and their ability to detect and repair miscommunication. We anticipated that children's EF and EK would be associated with more accurate ratings of their initial statements and with more effective communicative behaviour. That is, children with stronger EF and EK skills would be more accurate in determining the success of the trial, and the quality of their message based on the affective cues provided by the listener compared to those with lower EF and EK. We also anticipated that children's EF would be associated with higher quality of initial messages, and higher quality of repair statements, as has been shown in past research. Given that past research has shown that children learn from feedback over time (e.g. Bacso & Nilsen, 2017; Wardloe & Heyman, 2016), we also included the effect of trial in our models.

Method

Participants

Participants were 101 children (M_{age} = 5.10 years; *SD* = 0.53 years; 47 females) ranging in age from 4;0 to 5;11 recruited from elementary schools in a mid-sized Canadian city. The majority of children in the sample spoke English from birth (98%), and 19% of participants spoke another language other than English at home regularly.

The original sample was 109 children. However, data from participants who did not attempt to identify the target during the communication task on 3 or more trials (e.g., instead naming an alternate object to the target; n = 5) and from three participants who discontinued their participation partway through the communication task (n = 3) were not included in the analyses. Thus, a total of 8 participants were excluded from analyses.

Procedure and Materials

Participants completed tasks in the order presented below.

Communication task.

Task setup. Participants sat at a table across from a computer screen which showed the pre-recorded videos of another child of similar age (the listener, “Anne”) in seemingly live display. The experimenter sat beside the participant and discretely controlled which videos were played using a Bluetooth keyboard.

Warm-up procedures. Participants played a warm-up game with the listener on the computer screen. During this game, the experimenter told the participant that the child on the computer screen can hear them but cannot see them. The experimenter gave the child a picture of an object (a banana) and told the participant to give specific clues to the child on the computer screen in order for them to guess what the object is. For example, the experimenter asked participants what colour the object was and, after the participants said “yellow”, the experimenter played a pre-recorded video of the listener asking, “is it a lemon?”. After the participant provided three clues about the object to the child on the screen, a video was played of the child on the screen correctly guessing the object. This warm-up game demonstrated to participants that the listener on the computer screen could hear them and could respond to their messages but could not see the images in front of the participants.

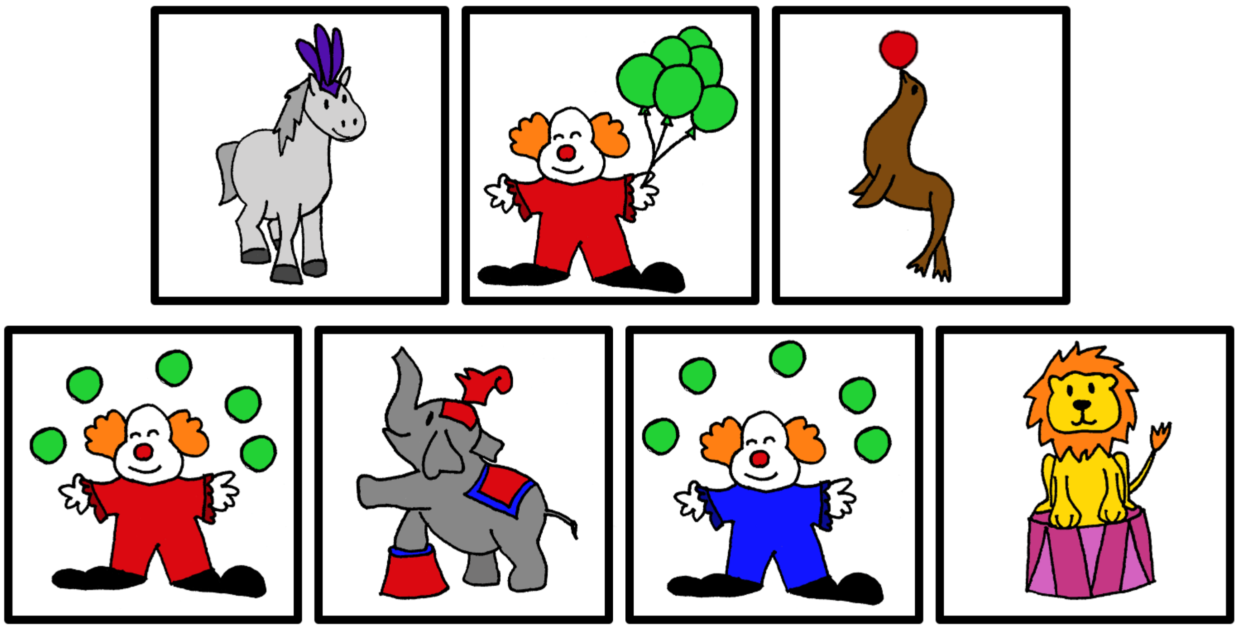
Practice trials. Children completed three practice trials. For practice trials, children were given a card depicting four boxes with different pictures on them (e.g. a flower, a sun, a cloud, and a tree). Children were told they would be helping the listener to find prizes. They were told that the listener has the same boxes in front of her as those shown on the child's card, and that the prizes were hidden in different boxes. The video panned across the boxes in front of the listener to show children that they were the same as those shown on their card, but in a different order. Participants were reminded that the listener can hear them but cannot see them so they would have to use their words to indicate which box the prize is hidden inside. The experimenter told the child, "It's in this one", and placed a token beside the picture of the target box. The child then told the listener which box the prize is inside. For practice trials, the boxes all had different pictures, so only the name of the picture was required to uniquely identify the target box (e.g. "the one with the sun on it"). Once the child provided an effective message, a video played of the listener picking up the target box and finding a candy inside.

Test trials. During test trials, children were again asked to describe which box the prize was located in for the listener. The experimenter showed the child which box the prize was hidden in by placing a token (which corresponded to boxes in front of the listener, albeit in a scrambled order). Contrary to practice trials, they were told that during test trials they would not be shown which box the listener chose. Children were also told that they could have another chance to tell the listener which box to choose if they wished to.

On each trial children were given a card depicting the images shown on the listener's boxes (see Figure 1). There were two different types of trials, which varied in the number of descriptors required to uniquely identify the target box. Stimuli for complex trials were designed such that children were likely to provide an ambiguous message on their first attempt, whereas stimuli on simple trials were designed such that children were likely to provide a uniquely identifying message on their first attempt. On *complex trials*, the image on the target box (e.g. a red clown juggling) was similar to those shown on several distractor boxes (e.g. a blue clown juggling and a red clown holding balloons), but was varied on two dimensions (i.e. colour, and associated object name). As such, to uniquely identify the target, two descriptors and the object name were required (e.g., “the one with the red clown juggling”). On *simple trials*, the image on the target box (e.g. a lion) was not similar to those shown on any of the other boxes (i.e. a clown, horse, etc.), and as such only the object name was required to uniquely identify the target (e.g., “the lion”). Thus, for complex trials, two descriptors were required to uniquely identify the target. On simple trials no descriptors were required. Children completed a total of 10 test trials (5 complex and 5 simple) in pseudorandom order.

Figure 1.

Example of Stimuli Used for the Communication Tasks.



Note. The child was told that the listener had boxes with the same cards printed on them. Each array had a total of 7 boxes, as shown here. For this set of stimuli, the complex target was the red clown juggling. The simple trial target was the lion.

Once the child provided their initial message, a video was shown of the listener reaching for a box, and her reaction to the box she opened. Children were not able to see which box was opened. If the child's message was ambiguous, the listener chose a box and looked sad[1]. If the child's message was uniquely identifying, the listener chose a box and looked happy. Regardless of the quality of the child's message on their first attempt, they were asked a series of questions about each trial by the experimenter after seeing the listener's facial expression:

I. Detection of miscommunication

1) *Do you think Anne found the prize this time?* **Success rating:** this determined the extent to which participants detected that miscommunication has occurred. Possible responses included "1" (*yes)* or "0" (*no).*
2) *How well do you think you described the box?* **Self rating**: this determined participants' ability to evaluate their own message. Children responded on scale from "-1" (*Not well),* "0" (*Okay)*, and "1" (*Well).*
3) *How well did Anne do at listening?* **Listener rating**: this question was used to determine whether participants attribute the success or failure of

[1] As a manipulation check, we showed children 4 screenshots of the listener's face (from the communication task) appearing happy and sad and asked them to identify whether the listener appeared happy or sad. Children were quite accurate in labeling the images with the correct emotion; 86% of children correctly identified the correct emotion on either 3 or 4 images. We also ran analyses excluding children who failed to correctly identify the correct emotion on 3 or 4 images. The pattern of results remained the same, with the exception that was no longer a significant interaction between executive functioning and facial cue in predicting speaker ratings (as presented in the Results section).

communication to the skills of the listener. They responded on a scale from "-1" (*Not well)*, "0" (*Okay)*, and "1" (*Well),*

II. Decision to repair

4) *Do you think you should try again?* [if yes] *Go ahead.* **Decision to repair:** this question was intended to determine whether children felt they should repair their initial message or not. If they choose to repair, the number of new descriptors and irrelevant descriptors provided was recorded.

III. Repair behaviour

Coding of communicative behaviour. Children's initial messages and their attempts to repair their messages were coded by a research assistant unaware of the research hypotheses. The responses were coded for the following: *Object name* (the name of the target object, e.g. "the boy*"*); *number of descriptors* (the number of descriptors provided during children's initial responses, e.g. "the boy in the *red shirt*, holding *ice cream"* would count as 2 descriptors); *irrelevant descriptors* (the number of descriptors provided during children's initial response or during repairs which do not include identifying information about the target, e.g. "the boy with *black shoes*", when all characters have black shoes); *new descriptors* (the number of new informative descriptors provided during a repair attempt that were not also included during the participants' initial response). Only number of descriptors and number of new descriptors were used for analyses.

A second research assistant coded the behaviours of 15 (15% of the total sample) randomly chosen participants to ensure reliability in coding. Interrater reliability for all coded responses was excellent (i.e., 99% agreement for the number of descriptors provided in initial statements and 96% agreement for the number of new descriptors following feedback).

Individual difference measures.

Executive functioning (EF). Children's EF was captured by a latent variable of children's performance on tasks of working memory, cognitive flexibility, and inhibitory control (described further in Results).

Inhibitory control. Children's inhibitory control was assessed using the naming and inhibition tasks from the NEPSY-II (Korkman & Kirk, 2007); though, due to time constraints, only one (rather than two) trials of each were used. Children completed a practice trial, followed by a naming trial, and an inhibition trial. During test trials, children were presented with a page showing 40 arrows arranged in rows and were asked to label the images as quickly as they could. On naming trial, children labelled shapes on a page (e.g. saying arrows are pointing "up" or "down"). After, children completed an inhibition trial, described similarly to the naming trials, however, children were asked to provide an incongruent label for the images (e.g. when they see an up arrow, they are asked to say "down"). This requires the child to inhibit their natural inclination to provide the correct label for the shape. The residual change score of the number of errors children made on the inhibition trial compared to the number of errors on the naming trial was used for analyses (max 40 errors for each trial). This provides a measure of children's inhibition skills while controlling for their naming skills. This change score was reflected (i.e., multiplied by -1), such that higher scores represented stronger inhibitory control skills (to be consistent with the other EF measures).

Cognitive flexibility. To assess children's cognitive flexibility, the Object Classification Task for Children (OCTC; Smidts, Jacobs, & Anderson, 2004) was used. During the task, children sort a series of objects which can be sorted based on size, function, or color (i.e. a small yellow plane, small red plane, large red plane, large red car, large yellow car, and small yellow car). Participants first sorted these objects into two groups in as many ways as they could. The experimenter then asked them to label their sorting criteria. Children received three points for each correct sort, and an additional point for accurately labelling their sorting criteria. If children were unable to sort the objects based on colour, size, and function, the experimenter sorted the objects into the groupings that were missed and asked the child to label the sort criteria. In this case, children received 2 points for correctly labelling each sort. If children were unable to correctly name the sort using this procedure, the experimenter then asked them to sort the objects based on the criteria that were missed. With this procedure, children received 1 point for each correct sort based on instructions by the experimenter. Total scores could range from 0 to 12.

Working memory. Children's working memory was assessed using the Digit Span subtest from the Wechsler Intelligence Scale for Children–Fourth Edition (Wechsler, 2003) as a measure of verbal working memory. Children first completed the digit span forwards task, in which they are asked to repeat back a series of digits read out loud by the experimenter. They then completed the digit span backwards task, where they were required to repeat a series of digits read by the experimenter in a backwards order. Children received a total score combining forwards and backwards tasks, out of a maximum possible score of 32.

Emotion knowledge (EK). To capture two aspects of children's EK (emotion labelling and emotion understanding), we combined the scores from the emotion labelling and emotion understanding tasks to create an EK composite.

Emotion labelling. Participants' emotion labelling was assessed using the Expressions subtest from the Assessment of Children's Emotion Skills (ACES; Schultz, Izard, & Bear, 2004). During the ACES, children saw 16 photographs of children displaying various emotions that they had to label as happy, sad, mad, or scared.

Emotion Understanding. Participants' emotion understanding was assessed using the Emotion Recognition Questionnaire (ERQ; Ribordy, Camras, Stefani, & Spaccarelli, 1988). This task assesses children's knowledge about different situations that elicit various emotions in others. During the task, children heard 12 (from the original 16) vignettes depicting a variety of situations and identified which emotion the main character was feeling out of three options provided by the experimenter.

Expressive vocabulary. Expressive vocabulary scores were assessed for use as a control measure using the Picture Naming task from the Wechsler Preschool and Primary Scale of Intelligence–Third Edition (Wechsler, 2002). For this task, children named a series of pictures that were presented to them. Children's raw scores (/24) were used for analyses.

Results

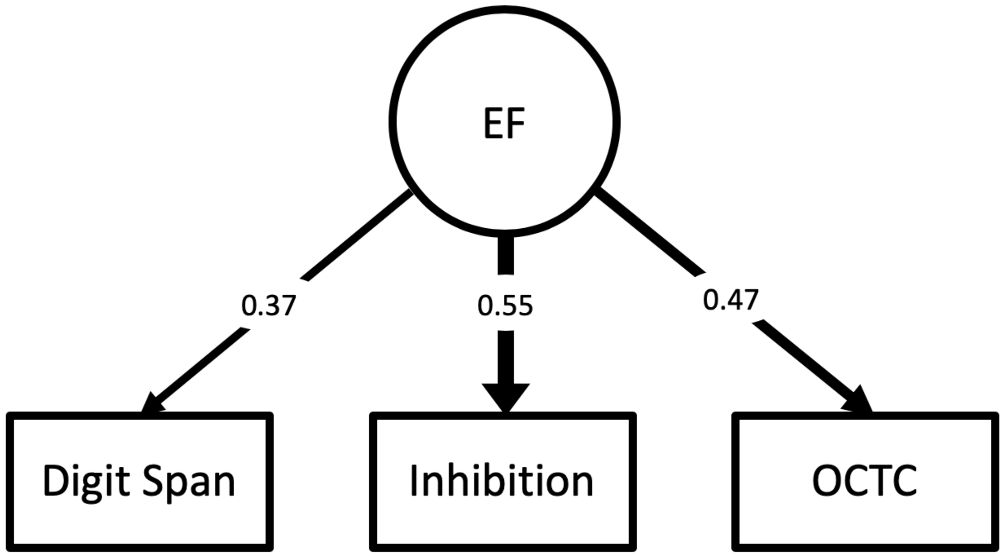
Preliminary Analyses

Two participants' scores were outliers for the number of errors on the inhibition task. Their scores on this measure were Winsorized to be within 3 standard deviations of the mean (as per Tabachnick & Fidell, 2007). For all continuous variables used in analyses, the standardized residuals of regression analyses showed normal distributions. Independent variables included in analyses also showed acceptable ranges for regression analyses (OCTC min = 4, max = 12; Digit span total score min = 3, max = 15; Inhibition total errors min = 0, max = 21; ACES min = 6, max = 16; ERQ min = 4, max = 12).

Model of Executive Functioning. A measurement model of an EF latent variable was created using children's digit span, OCTC scores, and reflected residual change scores on the inhibition task (see Figure 2). Analyses were conducted with the Lavaan package in R using maximum likelihood estimation (R Core Team, 2013). The latent EF factor was assigned a scale by using a unit loading identification constraint (i.e., the factor loading of the digit span total score was set to 1). To allow for enough degrees of freedom to assess model fit, the error terms of indicators were not allowed to covary within the model.

Figure 2.

Associations of the Latent Executive Functioning Variable with Executive Functioning Tasks.



Note. Beta (*β*) weights are shown in this model. The variables included were the digit span forwards and backwards total score, the reflected residual change score of the number of errors on the inhibition task, and the total score of the Object Classification Task for Children (OCTC). Associations between the inhibition task and OCTC task with the executive functioning latent variable were statistically significant ($p < .05$); however, as the digit span task was set as the scaling variable for the model its association with the executive functioning variable could not be determined through use of this model.

The chi-squared (χ^2) statistic and Root Mean Square Error of Approximation (RMSEA) were selected to provide measures of global fit. The χ^2 statistic suggested good model fit, $\chi^2(1, 101) = 1.91$, $p = 0.17$. The RMSEA score was slightly high, RMSEA = 0.095, $df = 1$, $p = 0.22$, falling just above the recommended cut-off value for an acceptable fit of 0.08 (Awang, 2012). The Comparative Fit Index (CFI) was selected to provide a measure of comparative fit and suggested acceptable model fit, *CFI* = 0.94 (Awang, 2012; Hair et al., 2006). Thus, the latent variable of EF was used for all further analyses.

Methods for Analyses

To examine whether children were sensitive to the listener's affective cues when evaluating their own messages and repairing miscommunication, as well as the skills that relate to their behaviour, we created several mixed models. Data were analyzed using multilevel modeling due to the hierarchical nature of our data (i.e. level one data would be the trials completed by participants, where level two data would be different participants). Multilevel modeling allows for residual components for each level of this hierarchy, thus controlling for idiosyncratic differences across trials within participants, and across participants. Multilevel models also are better equipped to deal with missing data compared to more traditional analyses. Models were created using the *lmer()* function and *glm()* function of the lme4 package (Bates, Maechler & Bolker, 2013) in R. Linear models (LMs) were used for numerical outcome variables; however, some of the outcome variables in this study were binary (e.g. success ratings and the decision to repair statements, with values of 1/yes or 0/no), and as such did not follow a normal distribution. Binary data was analyzed using generalized linear models (GLMs) which allow for dependent variables with binary data. All base models included random intercepts for participant as well as a fixed effect for trial. The fixed effect for trial was included as a control variable. Random slopes models were compared to these random intercept models to determine which provided the best fit. In most cases, random intercept models provided better fit than random slopes models, so these were used for analyses. When the random slopes models were used below, it is noted. To examine the effect of each variable, we added each variable to the model and compared the model fit to that of the base model. Models that did not result in significantly improved model fit were not examined further.

Throughout the analyses, the role of EF and EK were assessed separately since these variables were correlated ($r = 0.25$, $p < .01$), and thus, may limit our ability to capture the relationship that each EF and EK share with children's communicative behavior if included in the same model. Correlations between outcome variables when both sad and happy faces were presented can be found in Table 1 (excluding the number of new descriptors added during repairs because this only applies to sad face trials). Correlations between age, expressive vocabulary, and all outcome variables when a sad face was presented (i.e., repair needed) can be found in Table 2.

Models including covariates (age, gender, school attended, and expressive vocabulary) in addition to the variables of interest were compared with models which did not include covariates throughout the analyses in order to control for the influence of other possible characteristics of participants. In all cases, models including covariates did not demonstrate better fit than models without covariates, thus we report models without covariates.

Table 1

Means, standard deviations, and correlations with confidence intervals between communication task measures for all data.

Variable	*M*	*SD*	1	2	3	4
Detection of miscommunication						
1. Success rating	0.60	0.49				
2. Speaker rating	0.20	0.83	.57** [.51, .62]			
3. Listener rating	0.34	0.78	.46** [.39, .52]	.45** [.39, .52]		
4. Descriptors in initial messages	0.61	0.68	.34** [.23, .43]	.23** [.12, .34]	.11 [-.00, .22]	
5. Repair decision	0.61	0.49	-.22** [-.30, -.15]	-.16** [-.24, -.08]	-.12** [-.20, -.04]	-.02 [-.13, .09]

Note. *M* and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation. Success ratings refer to children's judgments on whether or not the trial was successful (i.e. whether or not she found the prize). This table includes data from all trials, i.e., both trials where the listener appears happy and trials where the listener appears sad. Note that the number of new descriptors provided during repairs were not included in this table, as the correlations between these and other variables are difficult to interpret when data from both happy and sad trials are included. The number of new descriptors provided during repairs can be seen in Table 2 * indicates $p < .05$. ** indicates $p < .01$

Table 2

Means, standard deviations, and correlations with confidence intervals for trials where a sad face was shown (i.e., repair needed)

Variable	*M*	*SD*	1	2	3	4	5	6	7
1. Age	5.10	0.53							
2. Picture Naming	19.24	3.55	-.06 [-.18, .06]						
3. EF	-0.02	0.71	.35** [.23, .45]	.06 [-.06, .18]					
4. EK	-0.05	1.03	.33** [.22, .44]	.24** [.13, .35]	.31** [.20, .42]				
Detection of miscommunication									
5. Success rating	0.31	0.46	-.24** [-.35, -.12]	-.10 [-.22, .02]	-.10 [-.22, .02]	-.15* [-.27, -.03]			
6. Speaker rating	-0.10	0.84	-.04 [-.16, .08]	.09 [-.03, .21]	.09 [-.03, .21]	.01 [-.11, .13]	.49** [.40, .58]		
7. Listener rating	0.12	0.81	-.09 [-.21, .04]	-.01 [-.14, .11]	.13* [.01, .25]	-.01 [-.13, .11]	.37** [.26, .47]	.41** [.31, .51]	
8. Repair decision	0.70	0.46	.17** [.05, .28]	.17** [.05, .29]	-.01 [-.13, .12]	.25** [.13, .36]	-.21** [-.32, -.09]	-.15* [-.27, -.03]	-.11 [-.23, .01]
9. Repairs: New descriptors	0.25	0.47	.16* [.02, .30]	.10 [-.05, .24]	.07 [-.08, .21]	.07 [-.08, .21]	-.16* [-.30, -.02]	.08 [-.06, .23]	-.09 [-.24, .05]

Note. M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation. This table includes only data from trials where a sad face was shown (i.e., the trial was unsuccessful). * indicates $p < .05$. ** indicates $p < .01$.

The results in the following sections use data from both the complex and simple trials. Since we designed the task to elicit miscommunication, such that children infrequently provided a uniquely identifying message on initial attempts during complex trials, some children only saw the listener look happy during a simple trial[2].

Detecting Miscommunication: Children's Judgments of Messages

Children's perceptions of their messages reflect the first stage of the repair process, in which children are required to first identify that their communication is in need of repair. We fit six mixed-effects models using the variables that provided information about children's perceptions of their own messages based on the nonverbal cues provided by the listener. There were two models for each dependent variable of success ratings, self ratings, and listener ratings. For each of these models, the base models were the same and included the within-participants predictor of trial, with random intercepts included for participants. Fixed effects for facial cue, EF and EK were added to the models. As mentioned, we examined EF and EK in separate models.

Success ratings. Recall that success ratings refer to children's judgements as to whether or not the listener found the prize on each trial. For success ratings, a one-way ANOVA revealed that the model including facial cue fit significantly better than the base model, $\chi^2(1, n = 593) = 198.33$, $p < 0.001$, suggesting a significant main effect of facial cue. This main effect is elaborated on in the analyses described below. Adding the latent variable of EF to the model that included facial expression, trial, and the random effect of participant did not improve model fit ($\chi^2(1, n =$

---

[2] To ensure that children were not solely using trial complexity as a cue about the quality of their message, we also re-analyzed the data using only data from the complex trials and found that the pattern of results was identical. This suggests that children were not only using the complexity of the stimuli as a cue to guide their ratings and decision to repair their messages.

587) = 0.01, *p* = .90), which suggests there was no main effect of EF. However, adding an interaction term of EF and facial cue significantly improved model fit, $\chi^2(2, n = 587) = 7.09$, $p = .03$. Thus, we used this model to assess the findings (See Table 3, Model 1).

Table 3
Details of the best-fitted mixed-effects models for success ratings.

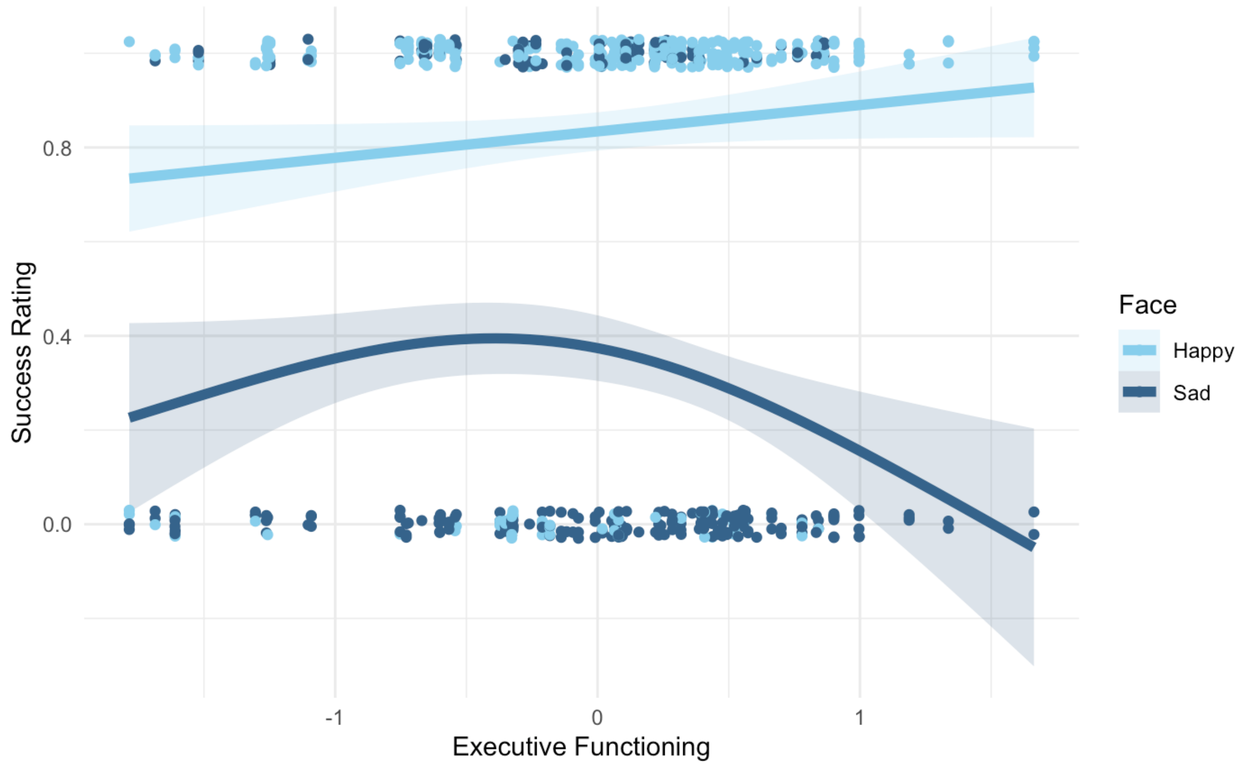
Model	Fixed Effects			
Model 1	Facial Cue + EF + Facial Cue*EF + Trial			
Executive Functioning	Predictors:			
		β	95% CI	*p*
	Facial Cue	-4.03	[-4.84, -3.22]	<.001
	EF	0.86	[-0.32, 2.03]	.15
	Facial Cue*EF	-1.06	[-1.84, -0.28]	.01
	Trial	0.61	[0.06, 1.15]	.03
Model 2	Facial Cue + EK + Facial Cue*EK + Trial			
Emotion knowledge	Predictors:			
		β	95% CI	*p*
	Facial Cue	-4.51	[-5.47, -3.56]	<.001
	EK	1.57	[0.19, 2.94]	.03
	Facial Cue*EK	-1.96	[-2.94, -0.97]	<.001
	Trial	0.77	[0.18, 1.35]	.01

Analyses revealed a significant main effect of facial cue, such that participants were more likely to rate the trial as successful following seeing the listener look happy (*M* = 0.84, *SE* = 0.02), compared to after seeing her look sad (*M* = 0.31, *SE* = 0.03). There was also a significant main effect of trial, where participants were more likely to rate the trial as successful on later trials, suggesting learning across trials. A significant interaction between facial cue and EF emerged and indicated that children with high EF were more likely to rate the trial as successful following seeing the listener look happy compared to those with weak EF skills (see Figure 3a). When the listener appeared sad, children with high EF were more likely to rate the trial as unsuccessful compared to those with weak EF. Therefore, children with high EF skills provided more accurate ratings of trial success based on the listener's facial cues.

Adding the composite variable of EK (i.e., instead of EF) to the model including facial cue did not result in improved model fit (p = .96). Adding the interaction term to the model significantly improved model fit, $\chi^2(2, n = 563) = 18.01$, $p < .001$. The model revealed a significant main effect of facial cue (see Table 3, Model 2), EK, and trial, as well as a significant interaction between facial cue and EK (see Figure 3b). This interaction suggested that children with higher EK skills were better able to detect whether or not the trial was successful based on the listener's facial expressions. For instance, children with higher EK were more likely to rate the trial as unsuccessful when the listener appeared sad, and successful when the listener appeared happy, compared to those with weak EK. Interestingly, a main effect of EK emerged within this model which indicated that children with higher EK were more likely to rate the trial as being successful than those with low EK, regardless of the facial expression of the listener.

Figure 3a.

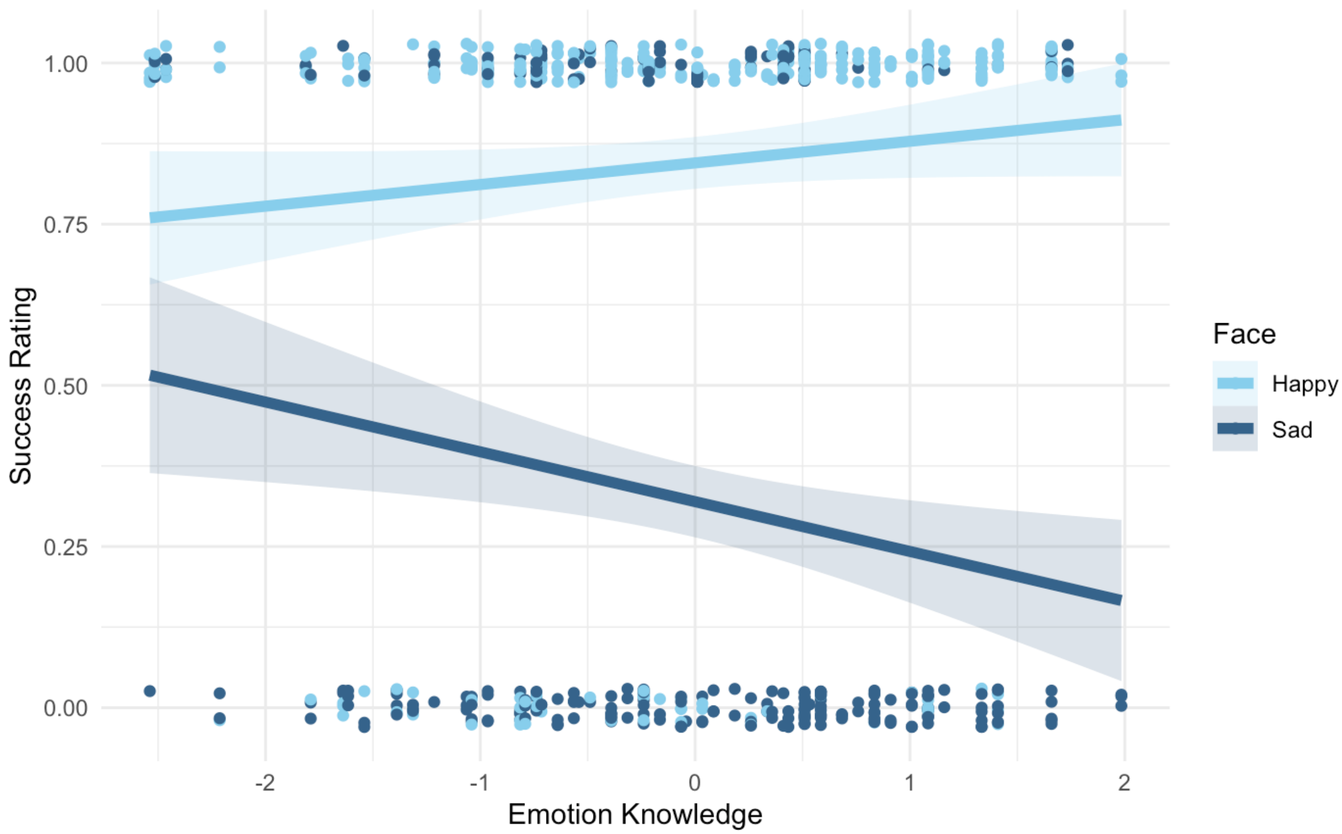
Children's Likelihood of Rating a Trial as Being Successful in Relation to the Listener's Facial Expression and to their Executive Functioning Skills.



Note. Points represent children's ratings on each trial as being successful (1) or unsuccessful (0), while lines represent the average ratings of children at each level of executive functioning skills. From this table you can see that children with stronger executive functioning are better at discriminating whether or not a trial was successful based on whether a sad or happy face was shown. Children with strong executive functioning appear better able to determine that the trial was unsuccessful when a sad face was shown.

Figure 3b.

Children's Likelihood of Rating a Trial as Being Successful in Relation to the Listener's Facial Expression and to their Emotion Knowledge Skills.



Note. Points represent children's ratings on each trial as being successful (1) or unsuccessful (0), while lines represent the average ratings of children at each level of emotion knowledge skills. From this table you can see that children with stronger emotion knowledge are better at discriminating whether or not a trial was successful based on whether a sad or happy face was shown.

Overall, for success ratings, we found that children were more likely to rate the trial as successful following seeing the listener look happy, compared to after seeing her look sad, with children who have higher skills in either EF or EK being more accurate at these determinations.

Self ratings. When examining children's ratings of their own skill, a one-way ANOVA revealed that the model including facial cue fit significantly better, $\chi^2(1, n = 593) = 62.91$, $p < .001$, suggesting a significant main effect of facial cue. We discuss this main effect within the models below.

Adding the latent variable of EF to the model including facial cue significantly improved model fit, $\chi^2(1, n = 564) = 7.91$, $p < .01$, which suggests there was a main effect of EF. Children with stronger EF were more likely to give their own performance as a speaker a higher rating than those with weaker EF, regardless of the facial expression of the listener. Adding an interaction term of EF and facial cue to this model significantly improved model fit, $\chi^2(1, n = 587) = 4.06$, $p = .04$. Within this model there was a significant main effect of facial cue (see Table 4, Model 1), such that participants rated their skills as a speaker higher on trials where the listener looked happy ($M = 0.43$, $SE = 0.041$), and lower after seeing her look sad ($M = -0.09$, $SE = 0.052$). There was no significant main effect of trial. A significant interaction between facial cue and EF indicated that children with high EF were more likely to rate their skills as speakers higher following seeing the listener looking happy compared to those with weak EF skills (see Figure 6). When the listener appeared sad, children with high EF were more likely to rate their skills as speakers on each trial lower compared to those with weaker EF. Therefore, children with high EF skills provided more accurate ratings of their skills as speakers based on the listener's facial cues.

Table 4

Details of the best-fitted mixed-effects models for speaker ratings.

Model	Fixed Effects			
Model 1	Facial Cue + EF + Facial Cue*EF + Trial			
Executive Functioning	Predictors:			
		β	95% CI	*p*
	Facial Cue	-0.31	[-0.38, -0.23]	<.001
	EF	0.23	[0.1, 0.36]	<.001
	Facial Cue*EF	-0.10	[-0.19, -0.001]	<.01
	Trial	0.01	[0.06, 0.08]	.75
Model 2	Facial Cue + EK + Facial Cue*EK + Trial			
Emotion knowledge	Predictors:			
		β	95% CI	*p*
	Facial Cue	-0.31	[-0.38, -0.24]	<.001
	EK	0.22	[0.09, 0.35]	0.001
	Facial Cue*EK	-0.15	[-0.25, -0.06]	.001
	Trial	0.01	[-0.06, 0.08]	.73

When EK was added to the model that included facial cue, it resulted in improved model fit, $\chi^2(1, n = 587) = 4.44$, $p = .04$. This indicates a significant main effect of EK, where children with stronger EK were more likely to rate their performance as a speaker higher than those with weaker EK. Adding the interaction term to the model significantly improved model fit, $\chi^2(1, n = 563) = 9.81$, $p = .002$. The model revealed a significant main effect of facial cue and EK (see Table 4, Model 2), as well as a significant interaction between facial cue and EK. This interaction suggested that children with higher EK skills were better at recognizing that they had provided a good description of the target when the listener appeared happy, compared to those with weak EK skills. When the listener's facial expression was sad, children all provided low self ratings, regardless of their EK skills.

Thus, we see that, again, children appear to be able to use the listener's facial cues to guide their perceptions of their communication, and that children with stronger EF and EK skills are more accurate in their judgments.

Listener Ratings. To further assess children's perceptions of the communicative exchange, we examined their ratings of the listener's skills on each trial. The analyses revealed comparable findings to the above variables and, thus, are only summarized here. Namely, while children generally rated the listener as being more skilled following seeing her look happy compared to after seeing her look sad, EF was associated with children making higher ratings of the listener overall. This suggests that children with stronger EF were more accurate in their ratings of the listener as the listener always followed their instructions perfectly (i.e., the listener always found the prize if the speaker gave an accurate description of the correct box). Children with higher EK skills rated the listener higher overall than those with weaker EK skills, with this effect most apparent when the listener appeared happy. When the listener's facial expression was sad, children's ratings were lower, with EK skills having less of an impact on children's ratings compared to when the listener appeared happy. The full description of these analyses can be found in a Supplementary File.

Children's Decisions to Repair and Repair Success

So as to provide some context in appreciating children's *repair* behavior, we first looked at how well children did during their initial statements. We focused on complex trials because these were the only trials in which descriptors were needed. In their initial descriptions, children uniquely identified the target on 11% of complex trials (they uniquely identified the target on 100% of simple trials). They provided a mean of 0.61 descriptors on each complex trial, which means that on average, 1.39 descriptors needed to be added during repairs to uniquely identify the target. Children's performance improved across complex trials (i.e., provided more descriptors during initial attempts, which indicates learning across trials), and older children performed better in general (See Supplementary File for full description of analyses).

To address our goal of examining children's repair behaviour, we looked at the next steps in the repair process (after detection), namely, their decision to repair, as well as their ability to provide necessary information (i.e., additional descriptors) to clarify misunderstood messages. The models for each analysis will be discussed in each section below.

Decision to repair. After seeing a sad face, children attempted to repair their messages on 70% of trials. We fit a mixed-effects regression model with a dependent variable of children's decision whether or not to repair their message on each trial. A one-way ANOVA comparing the random intercept and random slopes models indicated that the random slopes model fit significantly better, $\chi^2(2, n = 594) = 15.94$, $p < .001$, so the random slopes model was used for all analyses involving children's decision whether or not to repair their message.

Facial cue was added as a fixed effect to the model to assess the impact of the listener's facial expressions on children's decision to repair their initial message. A one-way ANOVA revealed that the model including facial cue fit significantly better, $\chi^2(1, n = 593) = 17.86$, $p < .001$, suggesting a significant main effect of facial cue. Children were more likely to repair their message following seeing the listener look sad ($M = 0.70$, $SE = 0.03$) than when the listener appeared happy ($M = 0.52$, $SE = 0.03$; see Table 5). A significant main effect of trial indicated that children were less likely to repair on each subsequent trial. This likely occurred because their initial descriptions significantly improved across trials, indicating learning across trials, so there was less need to repair messages.

Adding the latent variable of EF to the model including facial expression did not significantly improve model fit, which suggests there was no main effect of EF ($p = 0.99$). Adding an interaction term between EF and facial cue also did not improve model fit, suggesting there was no interaction between EF and facial cue ($p = 0.98$).

Table 5

Details of the best-fitted mixed-effects models for the decision to repair.

Fixed Effects			
Facial Cue + EK + Facial Cue*EK + Trial			
Significant Predictors:			
	β	95% CI	*p*
Facial Cue	1.02	[0.49, 1.54]	<.001
EK	0.34	[-0.74, 1.43]	.54
Facial Cue*EK	0.86	[0.18, 1.53]	.01
Trial	-1.30	[-1.83, -0.77]	<.001

Adding the composite variable of EK to the base model, did not result in improved model fit (p = .09). Adding the interaction term of EK and facial cue resulted in significantly better model fit, $\chi^2(2, n = 594) = 9.31$, $p = .01$. The model revealed a significant main effect of facial cue and trial, as well as a significant interaction between facial cue and EK (see Table 5). This interaction suggested that children with higher EK skills were better at recognizing that their message was in need of repair when the listener appeared sad. When the listener appeared happy, children's EK had less of an impact on their decision to repair.

Thus, with respect to children's decisions about repairing their messages, we see that they are able to use the listener's facial cue to guide whether or not they should attempt to repair their message. We see that EK (but not EF) is associated with children's ability to initiate a repair based on the listener's facial cues.

Repairs: Number of new descriptors. The analyses below include only trials where children attempted to repair their messages following seeing a sad facial expression. While children were using the listener's facial cue to guide repair attempts, their actual repairs were not very successful. That is, during repair attempts, children frequently repeated themselves (i.e., did so on 61% of attempts to repair initially ambiguous complex trials following seeing a sad facial expression). Children added new descriptors during repairs on 25% of trials, overall, children provided a mean of 0.25 new descriptors on each initially ambiguous trial. Children successfully repaired their messages (i.e., provided all descriptors needed) on 15% of initially ambiguous complex trials[3].

We found that participants provided fewer new descriptors with each subsequent trial (β = -0.18, 95% CI [-0.33, -0.26], $p < .001$). This is most likely due to participants providing more initial identifying descriptors on each trial, and thus being required to add fewer new descriptors during repairs. This provides further evidence of learning across trials.

Adding the latent variable of EF to the model including the random effect of participant did not significantly improve model fit, which suggests there was no main effect of EF (p = 0.39). Adding the composite variable of emotion understanding also did not result in improved model fit (p = .28). This suggests there was no main effect of emotion understanding on the number of descriptors added by children during repairs.

In sum, when examining children's communicative behaviour (i.e., decision to repair, new descriptors), we find that facial cues prompted more repair attempts, but the actual repair quality was relatively low within our task. This being said, children provided an increasing number of initial descriptors across trials, which suggests they were learning how to make their messages more effective throughout the task. Children with stronger EK were also more likely to attempt to repair their messages when the listener appeared sad.

---

[3] Altogether, including successful responses during initial statements and repairs, participants were able to uniquely identify the target on 26% of complex trials.

Discussion

This study examined children's abilities in detecting miscommunication, deciding to repair, and repairing miscommunication, in a context where only affective cues from a listener were provided. Further, children's EF and EK skills in relation to these steps were examined.

Detecting Miscommunication: Children's Judgments of Messages

To assess children's recognition of miscommunication and determination of who may be at fault for this miscommunication, we asked children to rate whether or not they thought each trial was successful, their performance as speakers on each trial, and the listener's performance on each trial.

Findings supported our hypothesis that children had accurate perceptions of their communication in a context where they were only provided with affective cues. That is, children's ratings across all areas indicated that they were able to detect whether or not the listener had found the prize on each trial based on the affective cues she provided. Specifically, children were more likely to rate the trial as being successful (i.e., the listener finding the prize) on trials where the listener appeared happy compared to trials where the listener looked sad. This finding is not surprising given past work which shows that even infants are able to understand that people's affective reactions should be consistent with the context (e.g., Chiarella & Poulin-Dubois, 2018). Importantly, in the present work, children were able to recognize their role in the success of each trial and rated their own performance as speakers as higher on trials where the listener appeared happy. This suggests that children show some degree of meta-communication, in that they are aware of their role in the success or failure of a communicative exchange. This being said, their listener ratings suggested that they adjusted their ratings of her skills based on the message success. Namely, children were likely to rate the listener as being less skilled following seeing her look sad compared to after seeing her look happy. Notably, the listener actually always responded in correspondence to what the child said (i.e., she always found the prize when children provided a uniquely identifying statement, and never found the prize following an ambiguous statement). So, really it was never the listener who was at fault when miscommunication occurred. Thus, children somewhat blamed the listener for their communicative errors (see Robinson & Robinson, 1983).

Together our results suggest that in a context where only affective cues are provided, preschool-age children show awareness of when miscommunication has occurred and, while they possess some insight as to their role in this, they do not take full responsibility for the miscommunication.

Communicative Repairs: Decision to repair and repair success

We anticipated that children would be able to use affective cues to guide their decision to repair or not repair their messages. Children demonstrated some recognition as to what they should do when provided with affective cues that suggest miscommunication occurred: they were more likely to attempt to repair their messages in response to seeing the listener appear sad, compared to seeing the listener look happy. Thus, in addition to verbal feedback (Bacso & Nilsen, 2017), preschool-age children can detect when to correct misunderstandings when provided only with nonverbal cues from their communicative partners. Wardlow and Heyman (2016) similarly found that 5- and 6-year-olds improved their messages across trials in response to nonverbal feedback indicating that the child's message was misunderstood. Notably, Wardlow and Heyman's study included the listener making an incorrect object choice, which prevents conclusions about children's use of facial expressions for repairs. The present findings suggest, that in addition to recognizing that the trial was unsuccessful when the listener appeared sad, children also understood that the listener required more information and sometimes attempted to provide such information.

However, even though children detected communicative ambiguity, they often failed to resolve this ambiguity. That is, children were only able to successfully repair (i.e., provide all the necessary descriptors) their initially ambiguous messages on 15% of trials in which a repair was attempted. Instead, children frequently repeated themselves (on 61% of trials). This is consistent with past literature indicating that children tend to repeat their messages in response to vague feedback (Anselmi et al., 1986; Nilsen & Mangal, 2012). Thus, children had limited success in repairing their messages. As with other studies using verbal feedback (e.g., Bacso & Nilsen, 2017) it is likely that repairs using nonverbal feedback are a skill which improves across the school-age years. This pattern is similar to that shown in other studies (albeit not ones specifically involving repairs) wherein children may recognize the need to provide additional information but fail to produce it. For instance, within a referential communication task, even when 4-year-old children looked at distractor objects (same object as target, but differing in one dimension), they fail to provide effective messages (i.e., those that uniquely identify the target) for their listener 83% of the time (Davies & Kreysa, 2017).

Our findings that children were able to detect when miscommunication occurred based on nonverbal cues and attempted to repair their messages support the assertion by Rabagliati and Robertson (2017) that children may require an “error signal” to detect whether or not they have avoided ambiguity to guide their learning in production. In Rabagliati and Robertson’s study, children (ages 4 to 5) demonstrated some evidence of self-monitoring their own messages (i.e. they looked at a distractor object after producing a message for a listener) but did not attempt to repair messages after this self-monitoring. In our work, children were able to use the affective cues of the listener as a signal to determine that their message required clarification, and then attempted to clarify their messages. As in other studies which provided more specific cues indicating ambiguity had occurred (Bacso & Nilsen, 2017; Matthews et al., 2007; 2012), children improved the clarity of their messages across trials. This suggests that they learned from miscommunication and from the error signal provided to adjust their production on subsequent trials. Thus, receiving affective cues and having the opportunity to repair their messages may be one way through which children's skills as speakers improve.

Role of Executive Functioning and Emotion Knowledge in Detecting Miscommunication and Repair

The second goal of the study was to investigate associations between EF and EK with children's perceptions of miscommunication and repairs.

With respect to children's perceptions, findings indicated that children with strong EF skills more accurately assessed the communicative scenario. Specifically, children with strong EF were more likely to rate the trial as unsuccessful and rate their own performance lower following seeing the listener appear sad. When the listener appeared happy, children with strong EF were more likely to rate the trial as being successful and rate their performance as speakers higher. Children's EF skills were also associated with their ratings of the listener: Children with higher EF skills generally rated the listener higher, which was a more accurate appraisal as the listener always responded appropriately to children's descriptions of the target. Thus, it seems they were less likely to blame the listener for miscommunication.

In the present work, a latent variable captured an underlying EF construct across the different tasks. A latent variable approach was used in order to reduce the influence of measurement error and task-specific variance (Kaushanskaya et al., 2017; Tabachnick & Fidell, 2013). Further, existing studies support the use of a latent variable to measure EF in both adults (Miyake et al., 2000), and young children (Fuhs et al., 2014; Willoughby et al., 2012). With this construct in mind, it appears that children's general ability to monitor and control thought and action facilitates more accurate perceptions of message and interlocutor success. To further appreciate how EF may be contributing to children's message evaluation, it is useful to consider the various aspects required in determining effectiveness of communicative utterances. For instance, children must be able to hold in mind their statements while simultaneously considering a listener's feedback (in the present study, nonverbal feedback), likely relying on working memory (see Bacso & Nilsen, 2017; Wardlow, 2013; Wardlow & Heyman, 2016). Cognitive flexibility may have supported children's flexible attendance to the various features of the target object in relation to the other objects (e.g. notice that a boy wearing a red shirt is also holding ice cream) and, use this information to determine message (in)effectiveness and the need for repair (see Bacso & Nilsen, 2017 & Gillis & Nilsen, 2014). Lastly, children were required to consider the perspective of their listener when assessing the communicative scenario (her knowledge and affective state), which may have required them to shift attention away from considering their own perspective, potentially drawing on inhibitory control skills (see Wardlow, 2013).

Similar to EF, children with stronger EK were better able to assess the communicative situation: They provided more accurate ratings for success of the trial, their own skills as speakers, and the skills of the listener based on the affective feedback provided. For example, children with strong EK skills were likely to rate the trial as being unsuccessful and rate their skills as speakers lower on trials where the listener appeared sad. Like children with strong EF, children with strong EK skills were also more likely than those with weaker EK skills to rate the listener higher in general, with this effect being strongest when the listener appeared happy. This is similar to the role of EK in listener ratings in that children with strong EK appeared to provide more accurate ratings of the listener.

EK likely aided children in understanding the meaning behind the listener's emotions (i.e., not just that the listener was sad, but that the reason was that she did not find the prize). This appreciation would facilitate children's ability to make accurate success ratings, self ratings, and listener ratings. Children with stronger EK were also more likely to attempt to repair their messages after seeing the listener appear sad, suggesting that they had a greater appreciation for the fact that the listener needed additional information. These findings are important as they demonstrate that EK adds to the skills which have been shown to contribute to communicative success, such as EF and theory of mind (see Nilsen & Fecica, 2011). Further, they suggest that EK skills allow for more effective interactions with peers, which may account for the findings shown previously between EK and social skills (Bassett, Denham, Mincic, & Graling, 2012).

We found that EK, but not EF was associated with children's decision to repair their messages. Neither skill was associated with children's success in repairing their messages. We had anticipated that EF would be associated with children's decision to repair and their repair success based on past research which found that children's EF was associated with both the quality of their initial messages and the quality of repairs, albeit following verbal feedback (Bacso & Nilsen, 2017). Thus, it may be the case that EF skills are important in contexts where children can use verbal feedback to guide their repair behaviours. However, the methodology of the present work may also account for different findings. In particular, in the present work, there was a time delay between children receiving the affective feedback and their opportunity to repair the message due to the questions asked. This may have interfered with the process of message repair that may have otherwise occurred, particularly for those children with high EF.

Future Directions and Limitations

While this work provides insight into children's evaluation of messages and repairs, there are some limitations to note. One design aspect worth noting is that the listener's affective reaction to the child's message was always consistent with the quality of the message. That is, the listener always appeared sad after the child provided an ineffective message, and always appeared happy after an effective message. This decision was made so that the task was naturalistic, and the listener could be viewed as providing reliable cues. However, it is possible that children were gauging the success or failure of their messages by purely reflecting on their own messages in the absence of any cues from the listener. We feel this explanation is unlikely given that repairs do not happen in the absence of any cues and past work has shown that children rely on their communicative partners for information as to whether their message was successful or not (Anselmi et al., 1986; Nilsen & Mangal, 2012; Wardlow & Heyman, 2016). Moreover, the relation between EK and children's perceptions of their communication suggests that the affective information was being utilized. Alternatively, it is also possible that children decided to repair their messages based solely on the affective cues provided by the listener, without reflecting on the content of their initial messages. Further research is needed to more precisely understand how children integrate the various information they could use when repairing messages. In addition, the task was designed to be difficult for children, as we wanted them to provide initial messages that were ineffective so that we could capture repair behaviour. However, given this, results should be interpreted within the demands of the task itself. A number of factors could influence the difficulty of referential communication tasks including the array size (our array included 7 items), the number of descriptors required for a uniquely identifying message, and the type of descriptors needed (i.e., some types of descriptors may be more readily produced by children, such as colour or size; Nadig & Sedivy, 2002). In addition, the type of feedback provided by the listener has been shown to influence children's ability to uniquely identify target objects (i.e., detailed feedback resulting in more successful repairs; Bacso & Nilsen, 2017). In the present study, children's task was also made more difficult by the fact that they were required to provide two descriptors *in addition to* the object name in order to uniquely identify the target. Lastly, other variables that were not included in this study likely have an impact on children's repair behaviour. For instance, children's level of shyness and/or their processing speed could have an impact on children's willingness/ability to repair messages.

We also found that the relations between EF and EK with children's communicative behavior were quite similar. This suggests there may be a shared element that accounts for the similar pattern of data for EF and EK. One possibility is that Theory of Mind (ToM) accounts for the associations between EF and EK and is related to children's communicative behavior. Past research supports this idea since children's executive functioning skills relate to their ToM skills (e.g, Carlson & Moses, 2001; Carlson, Moses & Breton, 2002) and children's ToM skills also predict their emotion knowledge skills (Hughes & Dunn, 2002; Seidenfeld et al., 2014).

Given our findings, future research could explore whether children are able to spontaneously and immediately repair their messages in response to affective, nonverbal feedback, and examine whether EF and EK are associated with this skill. This would expand upon the present work, in which we assessed whether children would repair their message after they were asked to reflect on the message. That is, as noted above, this delay may have influenced their natural inclination to repair messages and muted any associations between EF and EK in the quality of their repairs. Future work assessing their tendency to repair immediately following a nonverbal cue would shed light on the role of EF and EK in children's ability to repair their messages and would likely fit more closely to how children interact with others in the real world. Future work could also assess whether children's EK is associated with other aspects of their communicative skills, such as their comprehension of other's messages (wherein affect cues are embedded, e.g., San Juan et al., 2017; Berman, Chambers, & Graham, 2010).

Overall, the findings add to the literature exploring children's use of various cues from listeners to correct miscommunication (e.g., Bacso & Nilsen, 2017; Coon, Lipscomb, & Copple, 1982; Deutsch & Pechmann, 1982; Nilsen & Mangal, 2012; Uzundag & Küntay, 2018). In particular, the present study finds that, in a context where only affective cues are provided by a listener, 4- and 5-year-old children are able to accurately perceive the success of their messages, and attempt repairs appropriately (albeit with limited repair success). Further, within this context, children with better EF and EK demonstrated more accurate evaluations of message success and the role that they and the listener played in such communicative outcomes.

References

Anselmi, D., Tomasello, M., & Acunzo, M. (1986). Young children's responses to neutral and specific contingent queries. *Journal of Child Language*, *13*(1), 135-144. doi:10.1017/S0305000900000349

Awang, Z. (2012). Structural Equation Modeling Using Amos Graphic: UiTM Press.

Bacso, S. A., & Nilsen, E. S. (2017). What's That You're Saying? Children with Better Executive Functioning Produce and Repair Communication More

Bassett, H. H., Denham, S., Mincic, M., & Graling, K. (2012). The structure of preschoolers' emotion knowledge: Model equivalence and validity using a structural equation modeling approach. *Early Education & Development*, *23*(3), 259-279. doi:10.1080/10409289.2012.630825

Bates D, Maechler M, Bolker B. lme4.0: Linear mixed-effects models using s4 classes. 2013 Retrieved from http://lme4.r-forge.r-project.org/ (R package version 0.999999-3).

Berman, J. M., Chambers, C. G., & Graham, S. A. (2010). Preschoolers' appreciation of speaker vocal affect as a cue to referential intent. *Journal of Experimental Child Psychology*, *107*(2), 87-99. doi:10.1016/j.jecp.2010.04.012

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child development, 72*(4), 1032-1053. doi: 10.1111/1467-8624.00333

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development: An International Journal of Research and Practice, 11*(2), 73-92. doi: 10.1002/icd.298

Chiarella, S. S., & Poulin-Dubois, D. (2013). Cry babies and pollyannas: Infants can detect unjustified emotional reactions. *Infancy*, *18*, E81-E96. doi:10.1111/infa.12028

Coon, R. R., Lipscomb, T. J. & Copple, C. E. (1982). Effects of listener feedback on the messages of kindergarten children in a referential communication task. *Journal of Applied Developmental Psychology, 3,* 337-346. doi:10.1016/0193-3973(82)90006-5

Davies, C., & Kreysa, H. (2017). Looking at a contrast object before speaking boosts referential informativeness, but is not essential. *Acta psychologica*, *178*, 87-99. doi:10.1016/j.actpsy.2017.06.001

Denham, S. A. (1986). Social cognition, prosocial behavior, and emotion in preschoolers: Contextual validation. *Child development*, 194-201. doi:10.2307/1130651

Denham, S. A., Bassett, H. H., Way, E., Mincic, M., Zinsser, K., & Graling, K. (2012). Preschoolers' emotion knowledge: Self-regulatory foundations, and predictions of early school success. *Cognition & emotion, 26*(4), 667-679. doi: 10.1080/02699931.2011.602049

Denham, S. A. (1998). *Emotional development in young children.* Guilford Press.

Denham, S. A., & Couchoud, E. A. (1991). Social-Emotional Predictors of Preschoolers' Responses to Adult Negative Emotion. *Journal of Child Psychology and Psychiatry, 32*(4), 595-608. doi:10.1111/j.1469-7610.1991.tb00337.x

Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition*, *11*(2), 159-184. doi:10.1016/0010-0277(82)90024-5

Fabes, A., Eisenberg, N., Hanish, L. D., & Spinrad, T. L. (2001). Preschoolers' Spontaneous Emotion Vocabulary: Relations to Likability, *Early Education and Development, 12*(1), 11-27, doi:10.1207/s15566935eed1201_2

Fagan, M. K. (2008). Toddlers' persistence when communication fails: Response motivation and goal substitution. *First Language*, *28*(1), 55-69. doi:10.1177/0142723707083280

Gillis, R., & Nilsen, E. S. (2014). Cognitive flexibility supports preschoolers' detection of communicative ambiguity. *First Language*, *34*(1), 58-71. doi:10.1177/0142723714521839

Graham, S. A., San Juan, V., & Khu, M. (2017). Words are not enough: how preschoolers' integration of perspective and emotion informs their referential understanding. *Journal of child language*, *44*(3), 500-526. doi:10.1017/S0305000916000519

Grigoroglou, M., & Papafragou, A. (2019). Interactive contexts increase informativeness in children's referential communication. *Developmental psychology*. doi:10.1037/dev0000693

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). Multivariate Data Analysis. New Jersey: Pearson University Press

Hughes, C., & Dunn, J. (2002). 'When I say a naughty word'. A longitudinal study of young children's accounts of anger and sadness in themselves and close others. British Journal of Developmental Psychology, 20(4), 515-535. doi: 10.1348/026151002760390837

Kaushanskaya, M., Park, J. S., Gangopadhyay, I., Davidson, M. M., & Weismer, S. E. (2017). The relationship between executive functions and language abilities in children: A latent variables approach. *Journal of Speech, Language, and Hearing Research*, *60*(4), 912-923. doi: 10.1044/2016_JSLHR-L-15-0310

Khu, M., Chambers, C., & Graham, S. A. (2017). When You're Happy and I Know It: Four-Year-Olds' Emotional Perspective Taking During Online Language Comprehension. *Child development*. doi:10.1111/cdev.12855

Korkman, M., Kirk, U., & Kemp, S. (2007). *Nepsy-ii*. San Antonio, TX: Pearson.

Lloyd, P., Mann, S., & Peers, I. (1998). The growth of speaker and listener skills from five to eleven years. *First Language*, *18*(52), 081-103. doi:10.1177/014272379801805203

Martins, E. C., Osório, A., Veríssimo, M., & Martins, C. (2016). Emotion understanding in preschool children: The role of executive functions. *International Journal of Behavioral Development*, *40*(1), 1-10. doi: 10.1177/0165025414556096

Matthews, D., Biney, H., & Abbot-Smith, K. (2018). Individual differences in children's pragmatic ability: a review of associations with formal language, social cognition, and executive functions. *Language Learning and Development*, *14*(3), 186-223. doi:10.1080/15475441.2018.1455584

Matthews, D., Biney, H., & Abbot-Smith, K. (2018). Individual differences in children's pragmatic ability: a review of associations with formal language, social cognition, and executive functions. *Language Learning and Development*, *14*(3), 186-223. doi:10.1080/15475441.2018.1455584

Matthews, D., Butcher, J., Lieven, E., & Tomasello, M. (2012). Two- and four-year-olds learn to adapt referring expressions to context: effects of distracters and feedback on referential communication. *Topics in Cognitive Science*, *4*(2), 184-210. doi:10.1111/j.1756-8765.2012.01181.x

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive psychology*, *41*(1), 49-100. doi:10.1006/cogp.1999.0734

Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, *13*(4), 329-336. doi:10.1111/j.0956-7976.2002.00460.x

Nilsen, E. S., & Fecica, A. M. (2011). A model of communicative perspective-taking for typical and atypical populations of children. *Developmental Review*, *31*(1), 55–78. doi:10.1016/j.dr.2011.07.001

Nilsen, E. S., & Graham, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. Cognitive psychology, 58(2), 220-249. doi: 10.1016/j.cogpsych.2008.07.002

Nilsen, E. S., & Graham, S. A. (2012). The development of preschoolers' appreciation of communicative ambiguity. *Child Development, 83*(4), 1400-1415. doi: 10.1111/j.1467-8624.2012.01762.x

Nilsen, E. S., & Mangal, L. (2012). Which is important for preschoolers' production and repair of statements: What the listener knows or what the listener says?. *Journal of Child Language*, *39*(5), 1121-1134. doi:10.1017/S0305000911000432

O'Neill, D. K., & Topolovec, J. C. (2001). Two-year-old children's sensitivity to the referential (in)efficacy of their own pointing gestures. *Journal of Child Language*, *28*(01), 1-28.

R Core Team. R: a language and environment for statistical computing (Version 3.0.2) [Software]. Vienna, Austria: R Foundation for Statistical Computing; 2013. Retrieved from http://www.R-project.org.

Rabagliati, H., & Robertson, A. (2017). How do children learn to avoid referential ambiguity? Insights from eye-tracking. *Journal of Memory and Language, 94,* 15-27. doi:10.1016/j.jml.2016.09.007

Ribordy, S. C., Camras, L. A., Stefani, R., & Spaccarelli, S. (1988). Vignettes for emotion recognition research and affect education programs with children. *Journal of Clinical Child Psychology, 17*, 322–325. doi:10.1207/s15374424jccp1704_4

Robinson, E. J., & Robinson, W. P. (1983). Communication and metacommunication: Quality of children's instructions in relation to judgments about the adequacy of instructions and the locus of responsibility for communication failure. *Journal of Experimental Child Psychology*, *36*(2), 305-320.

San Juan, V., Chambers, C. G., Berman, J., Humphry, C., & Graham, S. A. (2017). The object of my desire: Five-year-olds rapidly reason about a speaker's desire during referential communication. *Journal of experimental child psychology*, *162*, 101-119. doi:10.1016/j.jecp.2017.05.003

Schultz, D., Izard, C. E., & Bear, G. (2004). Children's emotion processing: Relations to emotionality and aggression. *Development and psychopathology*, *16*(2), 371-387. doi:10.1017/S0954579404044566

Seidenfeld, A. M., Johnson, S. R., Cavadel, E. W., & Izard, C. E. (2014). Theory of mind predicts emotion knowledge development in head start children. *Early Education and Development, 25*(7), 933-948. doi: 10.1080/10409289.2014.883587

Smidts, D. P., Jacobs, R., & Anderson, V. (2004). The Object Classification Task for Children (OCTC): A measure of concept generation and mental flexibility in early childhood. *Developmental Neuropsychology*, *26*(1), 385-401. doi:10.1207/s15326942dn2601_2

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5). Boston, MA: Pearson.

Uzundag, B. A., & Küntay, A. C. (2018). Children's referential communication skills: The role of cognitive abilities and adult models of speech. *Journal of experimental child psychology*, *172*, 73-95. doi:10.1016/j.jecp.2018.02.009

Wardlow, L. (2013). Individual differences in speakers' perspective taking: the roles of executive control and working memory. *Psychonomic Bulletin Review*, *20*(4), 766 – 772. doi:10.3758/s13423-013-0396-1

Wardlow, L., & Heyman, G. D. (2016). The roles of feedback and working memory in children's reference production. *Journal of Experimental Child Psychology, 150*, 180-193. doi:10.1016/j.jecp.2016.05.016

Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence* (3rd ed.) San Antonio, TX: Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children* (4th ed.) San Antonio, TX: Psychological Corporation.

Wechsler, D. (2009). *Wechsler Individual Achievement Test* (3rd ed.). San Antonio, TX: Psychological Corporation.

Wilcox, M. J., & Webster, E. J. (1980). Early discourse behavior: An analysis of children's responses to listener feedback. *Child Development*, 1120-1125. doi:10.2307/1129

Willoughby, M. T., Wirth, R. J., & Blair, C. B. (2012). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological assessment, 24*(2), 418. doi: 10.1037/a002577

Appendix

Items included in the communication task. Note that items were presented in one of three possible orders for counterbalancing and the same array of boxes was used for both a simple and a complex trial. Key descriptors for complex trials are bolded. Key distractor items for complex trials are in italics.

Target object on complex trial	Target object on simple trial	Distractors
Red clown **juggling** or **with balls**	Lion	*Blue clown juggling, red clown holding balloons*, horse with feathers on its head, seal with a ball, elephant on a platform, lion on a platform
Boy **wearing red holding an ice cream**	Dog	*Boy wearing green holding an ice cream, boy wearing red holding a drink*, girl on a bench, man holding ice cream, dog with a bone, cat with a mouse
Brown monkey **holding a banana**	Tiger	*Black monkey holding a banana, brown monkey holding a flower*, tiger with a leaf, hippo with a flower, giraffe with a ball, toucan with bananas