Children's language abilities predict success in remote communication contexts

Karla K. McGregor
Ronald Pomper
Nichole Eden
Timothy Arbisi-Kelm
Nancy Ohlmann
Shivani Gajre
Erin Smolak
Boys Town National Research Hospital, USA

Abstract: Remote communicative contexts are part of everyday social, familial, and academic interactions for the modern child**.** We investigated the ability of second-graders to engage in remote discourse, and we determined whether language ability, theory of mind, and shy temperament predicted their success. Fifty 7-to-9-year-old monolingual English speakers with a wide range of language abilities participated in standardized testing and an expository discourse task in which they taught two adults to solve the Tower of London, one in an audiovisual condition to simulate video chat and a second in an audio-only condition to simulate phone communication. The discourse was scored with a rubric of 15 items deemed relevant to the explanation. Children included 27% to 87% of the items, with more items communicated via gesture than spoken word in both conditions. Gesture scores and spoken scores were highly correlated. Children specified more rubric items overall in the audio condition and more rubric items in the spoken modality when in the audio condition than the audiovisual condition. Performance in both conditions was positively associated with scores on independent measures of language ability. There was no relationship between performance and theory of mind, shy temperament, ability to solve the Tower of London, age, or sex. We conclude that 7-to-9-year-olds adjust the modality and content of their message to suit their remote partner's needs, but their success in remote discourse contexts varies significantly from individual to individual. Children with below-average language skills are at risk for functional impairments in remote communication.

Keywords: remote communication; pragmatics; gesture; phone; video chat

Corresponding author(s): Karla McGregor, Center for Childhood Deafness, Language & Learning, Boys Town National Research Hospital, 425 N. 30th St., Omaha, NE, 68131, USA. Email: karla.mcgregor@boystown.org.

ORCID ID(s): https://orcid.org/0000-0003-0612-0057

Citation: McGregor, K.K., Pomper, R., Eden, N., Arbisi-Kelm, T., Ohlmann, N., Gajre, S., & Smolak, E. (2021). Children's language abilities predict success in remote communication contexts. *Language Development Research,* 1(1), p 245-282. DOI: 10.34842/8jgf-r802

Introduction

Children modify their communication according to audience and context (Akhtar et al., 1996; Mori & Cigala, 2016; Nadig & Sedivy, 2002; Nilsen & Fecica, 2011; Shatz & Gelman, 1973). For example, four-year-olds speak in shorter sentences and use more utterances to draw the listener's attention when addressing toddlers than peers or adults (Shatz & Gelman, 1973). However, this early emerging skill must be honed over developmental time as children encounter new contexts. In this paper, we considered remote communication, a context that is prominent in the everyday lives of children (McClure et al., 2015; MPFS, 2018) and particularly so since the first quarter of 2020 when many families were quarantined in response to the COVID19 pandemic (Koeze & Popper, 2020). We investigated second-graders' ability to engage in discourse via an audio-only channel to simulate phone communication and via an audiovisual channel to simulate video chat, and we determined whether language ability, theory of mind, and shy temperament predicted their success.

Children's Phone and Video Communications

From birth, children have access to various technologies with which they learn to share meaning with others (Erstad et al., 2020). Children's phone talk has long been recognized as a developmental step between the 'here and now' audible and instantaneous exchange characteristic of face-to-face talk and the decontextualized nature of written communication (Cameron & Lee, 1997; Gillen, 2002). Children talk on the phone long before they can read and write yet, to do so successfully, they must heed the needs of listeners who do not share their immediate context, just as the successful writer must.

Children's phone communication skills emerge early but follow a protracted developmental course. Take, for example, a study by Cameron and Lee (1997). They had three to eight-year-olds solve the Tower of London (Shallice, 1982), a task in which the child must move beads, one at a time, onto pegs to match an array demonstrated by the examiner in as few moves as possible. Afterward, they asked the children to explain the task to listeners face-to-face or on the phone. The older children gave more complete directions than the younger children. Children of all ages provided more detail and used more specific vocabulary while speaking on the phone than in person (Cameron & Lee, 1997). However, their communication was not necessarily worse when in person. The children used frequent visual checking of the listener's performance, presumably to adapt the directions to the listener's needs; what they accomplished with shared visual contexts when face-to-face, they accomplished with more extensive, detailed instructions when on the phone.

An early ability to adjust to a remote listener's needs is also evident in narrative discourse. For example, Pinto et al. (2016) found that 5-to7-year-olds make more mention of characters' mental states when narrating a pictured event during phone talk than when face-to-face, presumably because they realize that the listener cannot intuit the characters' mental states without seeing the pictures.

Children's communication on video-chat platforms has been less extensively researched than their phone talk. However, here too, we find very young children demonstrating some success. For example, McClure et al. (2018) observed families with babies ages 6 to 24 months as they engaged in video chats with the babies' grandparents. They were particularly interested in times when the baby initiated joint attention with the grandparent across the screen by, for example, showing a toy. Whereas only 8% of the babies under 15 months ever did so, 46% of the babies older than 15 months engaged their grandparents in this way. This extremely early performance reflects an essential difference between video chat and phone talk, shared visual context.

In complicated communicative exchanges, even adults may benefit from shared visual context. For example, Veinott et al. (1999) asked native and non-native English speakers to explain a mapped route to a listener who joined them via video or audio-only. Although the native-speaking pairs were equally successful in the video and audio conditions, the non-native speakers benefitted from the video context. Specifically, they were faster and more accurate at communicating the route in the video condition. This success was engendered by more talk devoted to instruction, more checks on mutual understanding, and more frequent gesturing. However, gesture was not considered in the Cameron and Lee (1997) comparison of children's phone and in-person communication; therefore, we do not know the extent to which children's communication might be enhanced by gesture in remote communication contexts that involve shared visual access.

The Development of Discourse in the Gesture and Speech Modalities

Gesture and speech are integrated systems of communication. They work together to convey meaning and affect, and they do so in ways that benefit both the listener and the speaker. When compared to speech alone, speech plus gesture reliably enhance listeners' comprehension. The benefits are greater for gestures that convey actions (e.g., how to do something) rather than abstract concepts (e.g., how something feels) and for gestures that complement rather than reproduce the meanings conveyed by words (Hostetter, 2011). In addition, children benefit from gestured input more than adults (Hostetter, 2011). Not surprisingly then, children tend to learn more readily from teachers who frequently gesture in ways that accurately convey new concepts than from teachers who do not (Alibali et al., 2013; Ovendale et al., 2018).

For the speaker, gesture assists with memory and planning the information to be conveyed (Alibali et al., 2000). Moreover, gesture supports thought. Under high cognitive load, like solving a problem, counting, or reasoning out load, children's gestures reflect their understanding (Alibali & Goldin-Meadow, 1993; Broaders, et al., 2007; Church & Goldin-Meadow, 1986; Ehrlich et al., 2006; Garber & Goldin-Meadow, 2002; Göksun et al., 2010; Pine, et al., 2004). When on the cusp of understanding, children often convey a more accurate grasp with their gestures than with their words, and this gesture-speech mismatch indicates learning readiness (Alibali & Goldin-Meadow, 1993).

Children communicate with gestures before they speak their first word (Capone & McGregor, 2004). However, gesture, like spoken language, continues to develop well into the school years. Alamillo et al. (2013) observed six- and 10-year olds during monologic narrative and dialogic explanation tasks. The older children used not only more complex spoken language but also more complex gestures. Both groups of children used more gestures during dialogue than monologue, suggesting some awareness of their partner's needs during the dialogic exchange.

Roth (2002) argues that gesture provides a stepping stone along the path of discourse development and in the expression of new knowledge during discourse. He observed tenth-grade students over multiple physics lessons as they conducted hands-on experiments and then explained their findings. The result was a robust developmental progression. First, the students spoke as they repeated their actions on the objects used in the experiments and, soon after, as they simulated the actions on other arbitrary objects. Gradually, they supplemented their spoken explanations with gestures produced without the support of objects. In these early attempts, the students tended to be more accurate in their actions on objects and their gestures in general than in their spoken explanations. Finally, the students arrived at a mature understanding of the physics problems they expressed in written or spoken words. At this mastery stage, spoken language and gesture continued to co-occur, although the gesture frequency was lower than in the more novice stages. In this way, the developmental course of these much older children faced with a new and complex task recapitulated the early communicative development of toddlers. Both demonstrate a progression from actions on objects to gesture to speech while never abandoning any of these highly functional modalities (Capone & McGregor, 2004).

The Contribution of Language Ability, Theory of Mind, and Shy Temperament to Successful Remote Communication

The success of any communicative interaction and the extent to which speakers can adapt their gestured and spoken communication to contextual demands will depend, in part, on the language abilities of the interlocutors. When communicating on the phone, the lack of shared visual context means that the speaker must provide information via words rather than gestures and that the words and their organization must be specific enough to enable comprehension. In these decontextualized exchanges, language *becomes* the context, and thus the communication partners must have the lexical, syntactic, and discourse skills necessary for creating clarity and common ground (Uccelli et al., 2019). When communicating via video chat, the shared visual context lessens these demands. However, challenges remain relative to face-to-face communication where the extent of the shared visual context is greater, the signal-to-noise ratio is higher, and physical interaction is viable.

Another factor influencing the success of communicative exchange is the interlocutors' ability to engage in theory of mind (Hughes & Leekham, 2004; Miller, 2006). Theory of mind refers to perceiving one's own and others' emotions, beliefs, desires, thoughts, and knowledge systems. Theory of mind develops from early childhood well into adolescence (Dorval et al., 1984). A speaker with a mature theory of mind will recognize the listener's need for more or less information.

Nevertheless, remote communication may present some challenges to perspective-taking, even for those with a strong theory of mind. Keeping track of the listener's perspective without a shared visual context, as during phone conversations, may impose a high memory load (Zhao et al., 2018). Moreover, primate work demonstrates that face, hand, and body movements provide essential cues to accurate social perception (Allison et al., 2000). Adult humans will even infer mental states from the movement of animated abstract shapes (Castelli et al., 2002). During remote communication, movements that cue the extent of the listener's understanding are limited (in video chat) or missing (in phone interactions).

The temperament of the interlocutors may influence the success of communication as well. Temperament is a stable trait that is highly heritable (Buss & Plomin, 1984; see Henderson & Wachs, 2007, for a review). Children who possess a shy temperament are inclined to withdraw from social interaction, particularly in unfamiliar social situations (Schmidt & Tasker, 2000). Shy children are reticent to talk; they talk less, make fewer spontaneous remarks, and are more likely to be unresponsive to strangers than their outgoing peers (Prior et al., 2000; Smith Watts et al., 2014). Thus, the high verbal demands of remote communication contexts may be especially challenging for shy children. That said, remote communication may be more comfortable than face-to-face communication for some shy children, in which case they may be

less reticent than documented in previous (in-person) research.

Although they are separate constructs, language ability, theory of mind, and shy temperament are interrelated. The relation between language and theory of mind is well studied. A meta-analysis of 104 studies with English-speaking children below age seven revealed a significant positive relationship between language and theory of mind abilities (Milligan et al., 2007). For example, vocabulary and grammar skills at two (Farrar & Maag, 2002) and grammar skills at three (Astington & Jenkins, 1999) predict theory of mind skills two years later. Language and shyness are also related. Specifically, shy children tend to score lower on formal tests of pragmatics (Copelan & Weeks, 2009), receptive vocabulary, and phonemic awareness (Spere et al., 2004) than their more outgoing peers, although not so low to be of clinical concern.

In contrast, the relation between shyness and theory of mind is not fully understood. Some investigators find that shy children perform more poorly on theory of mind tests than outgoing children (Banerjee & Henderson, 2001; DeRosnay et al., 2014; Walker, 2005), leading them to hypothesize that a shy temperament limits their social-communicative interactions, and thus their opportunities to learn more about reading another's mind. On the other hand, others report that shy children demonstrate an advantage on theory of mind tests (Mink et al., 2014; Wellman et al., 2011), leading them to hypothesize that shy children sharpen their theory of mind by observing others' social-communicative interactions.

The Current Study

In the current study, we examined the expository discourse of second graders in two remote communication contexts, one that simulated phone communication by providing an audio channel only and the other that simulated video chat communication by including both audio and visual channels. We were particularly interested in second graders because they are in the throes of language and theory of mind development and, by second grade, their temperament is highly stable (Neppl et al., 2010). Moreover, they are still early in their formal reading and writing instruction years, a time when individual differences in bridging the fully contextualized nature of face-to-face talk and the fully decontextualized nature of formal writing are likely to be high.

Expository discourse typically involves more complex syntax and more specific or sophisticated vocabulary choices than conversational discourse. We selected an expository discourse task because it is more likely than conversational exchange to reveal individual differences between children. For example, adolescents with language impairments do not differ significantly from their typical age-mates in conversation but, during exposition, they tend to use shorter, less syntactically complex sentences (Nippold et al., 2008). Expository discourse is also high in ecological validity given that

mastery of expository discourse is recommended as an instructional goal in the academic curriculum (CCSS; Common Core State Standards Initiative, 2015), and it is the

type of discourse required for everyday communicative goals such as giving directions or explaining the rules of a game (Lundine & McCauley, 2016).

Like Cameron and Lee (1997), we used the Tower of London to elicit the expository discourse. The task requires problem-solving and planning (visualizing several moves ahead) and other aspects of executive function such as attention, memory, and inhibition. Language, either internalized or externalized, is helpful for scaffolding performance on the Tower of London. Perhaps, as a result, children with developmental language disorder tend to perform poorly on the task (Larson et al., 2019; Marton, 2008; Roello et al., 2015, and verbal suppression impairs performance in children with and without language disorder (Lidstone et al., 2012). After completing the Tower of London, we asked the children to explain the game to two naive adults, one who was not present but could hear them (audio condition) and one who was not present but could hear and see them (audiovisual condition).

Questions and Hypotheses

We preregistered the study (McGregor et al., 2019, available at this link: OSF Registries | Children's Voabulary Project; Remote Communication) as a comparison between second graders with and without developmental language disorder, a prevalent neurodevelopmental condition characterized by limitations in language learning, comprehension, and use. Unfortunately, because of the COVID-19 pandemic, we were forced to close the study before recruiting enough participants with developmental language disorder. Nevertheless, we had an excellent distribution of language abilities represented in the sample and adequate power to investigate language ability as a continuous predictor of remote communication performance. Thus, we modified our predictions to be:

Children would provide more complete directions in the audiovisual than audio condition because they would more frequently supplement their verbal message with gestures in the latter than the former.

There would be individual differences across children such that those who have more robust vocabularies, who are less shy, and who have a more highly developed theory of mind would be more successful on the task than those who scored lower in these domains. We also determined the effect of age, sex, and success on the Tower of London itself.

Finally, we took this opportunity to explore the relationships between vocabulary,

theory of mind, and shyness to address incomplete or conflicting reports in the literature.

Methods

Participants

The project was conducted in compliance with protocols approved by the Internal Review Board of Boys Town National Research Hospital to ensure the protection of human subjects. Participants were 50 second graders (29 girls), ages 7 to 9 years (median = 100 months, min-max = 88 to 109). Two additional children were tested but are not included here, one because of a subsequently diagnosed seizure disorder and the other because of attrition.

All participants were monolingual English speakers from Iowa or Nebraska in the United States recruited from a larger longitudinal study of language development (Research Registry 3425, 2017). According to parents' reports of ethnicity, one participant was Hispanic or Latino, 41 were neither Hispanic nor Latino. Eight parents did not report ethnicity. According to parents' reports of race, one participant was African American, 43 were Caucasian, and six were more than one race.

The children presented with a range of spoken language abilities, with standard scores from 72 to 127 on the *Test of Narrative Language-second edition* (TNL-2, Gillam & Pearson, 2017). Eleven were receiving special support for language in or outside of school. To ensure that neither intellectual disability nor hearing loss contributed to variability in task performance, we limited enrollment to participants who earned a perceptual index score of 70 or higher on the *Wechsler Abbreviated Scale of Intelligence* (Wechsler, 1999) and passed a pure tone audiometric screening.

Procedure

We administered standardized tests to determine the abilities that predict expository discourse performance, and then we administered the expository discourse task itself. Data collection occurred over two or three sessions scheduled within two weeks.

NIH Toolbox Picture Vocabulary Test

The *NIH Toolbox Picture Vocabulary Test* (Gershon et al., 2013) measures receptive single-word vocabulary. The participant is instructed to touch the image in a 4-alternative forced-choice array that they believe is most closely associated with the word they heard. The difficulty level of each trial is automatically adjusted by the software program, contingent on the participant's previous response's accuracy. Raw scores on this task were converted to normally distributed standard scores (scaled scores) that

were not age-corrected. Specifically, the raw scores were ranked and then transformed to create a standard normal distribution, which was then re-scaled to have a mean of 10 and a standard deviation of 3.

Temperament in Middle Childhood Questionnaire (TMCQ)

The TMCQ (Simonds & Rothbart, 2006) taps caregiver judgments of the emotional temperament of children between 7 and 10 years of age. Although the TMCQ measures a broad range of temperament traits, we only used the items that tapped shyness. The temperament dimension "shyness" is operationally defined in the TMCQ as "slow or inhibited approach in situations involving novelty uncertainty." The shyness score was calculated by summing the ratings for the five questions categorized within the shyness temperament dimension.

Theory of Mind Inventory-2 (ToMI-2)

On the ToMI-2 (Hutchins & Prelock, 2016), caregivers mark along a 5-point continuum ranging from “Definitely Not” to “Definitely” to describe the most likely way their child would mentalize in 60 different scenarios. The questionnaire is scored by plotting each response along a 20-centimeter scale and rounding the score to the nearest tenth. In the present analysis, these scores were then summed and divided by 60 to derive a mean score ranging from 1 to 20.

Tower of London (ToL)

The Tower of London Drexel University -2nd Edition (Culbertson & Zillmer, 2005) measures problem-solving and planning. In the ToL, both the children and the examiner use boards containing three pegs of varied sizes holding one to three colored beads. The examiner demonstrates beads stacked in 10 different arrangements on wooden pegs. The goal is for the child to move their beads one at a time and in as few moves as possible from a start position to match the examiner’s array.

We administered the task according to the directions in the test manual. The critical directions given to the children were: 1) the two pegboards must be alike, 2) as few moves as possible must be used to copy the design on the examiner's board, 3) no peg may contain more beads than it can hold, and 4) only one bead at a time can be moved, in other words, two or more beads cannot be taken off the board at one time. The score we derived was the number of moves needed to solve each. Four of the participants skipped one (N = 3) or 5 (N = 1) items on the ToL due to experimenter error or participant fatigue; therefore, we transformed scores by dividing the total number of extra moves (total moves – minimum moves) by the minimum number of moves for each child. A child who completed each arrangement in the minimum number of

moves would have a proportion of 0, whereas a child who completed each arrangement with twice the number of the minimum possible moves would have a proportion of 1.

Expository Discourse

After completing the ToL, we asked the child to explain its procedures and five example problems to two naive adults who were not present, one who could hear them (audio condition), and one who could hear and see them (audiovisual condition). By simulating phone and video chat rather than engaging the children in these actual contexts, we ensured that the performance was the child's own, not the result of more or less scaffolding from a communicative partner. The exact instructions are included in the Appendix. All children participated in both conditions with order counterbalanced across participants.

In the audio condition, we showed the child a photograph of an unfamiliar adult woman. We said that she did not have access to video technology but would hear the child's instructions when she called later. A phone was included in the photograph. The children were then asked to explain to their listeners what the game looks like, what the rules are, and exactly how to play.

In the audiovisual condition, we showed the child a photograph of a second unfamiliar adult woman. We said that she had access to a computer (a computer was included in the photo), so she would hear and see the child's instructions when she logged in later.

The entirety of the data collection session was video recorded via a laptop camera for later scoring. For the audiovisual condition only, we also recorded with a camera on a tripod to illustrate more clearly to the child that their remote partner would be able to not only hear them but also see them. An example of a child participating in the aduio and audiovisual conditions is available at OSF | Example of child completing the discourse task.

Discourse Scoring

A 15-item rubric was created to capture the pragmatic and semantic content of each child's discourse (see Appendix). The child could receive one point for each of the first 11 items in the rubric, regardless of whether they expressed that item in gesture, spoken words, or both. Gestures could be representational (e.g., making a circle shape to indicate a bead), deictic (e.g., pointing or showing), or demonstration (e.g., moving the bead from one peg to the next). The items scored were: item 1, introducing the discourse (e.g., saying or gesturing hello); items 2 – 5, explaining each of four rules well enough for a naive listener to apply the rule successfully; items 6 – 10, explaining

each of five trials well enough for a naive listener to complete the trial successfully; and item 11, closing the discourse (e.g., saying or gesturing goodbye, we're done). In addition, we were interested in the children's use of vocabulary deemed essential to the explanation. Thus, they could also earn one point for each of four spoken vocabulary items. These were at least one mention of 1) the bead/ball, 2) the peg/stick/stand, 3) the location (e.g., here, long peg), and 4) the sequential order (e.g., next, last). A second coder independently scored 22% of the discourse samples. The point-to-point agreement was 91.94%.

When using the overall score as a dependent variable, the highest possible score was 15: 11 spoken and/or gestured items + 4 spoken words. Because some children skipped items due to experimenter error or participant fatigue (in the audio condition, 10 participants skipped one item and one participant skipped two; in the audiovisual condition, six participants skipped one item), we transformed scores into proportions (points received/total maximum points possible given items administered) and applied a logit transformation.

Statistical Analysis

The preregistered data analysis plan was to use a linear mixed-effects model with the explanation score as the dependent variable and independent variables including a fixed within-subjects factor of condition (audio or audiovisual) and between-subjects effects of diagnosis (DLD or TD), sex (M, F), age, diagnosis x sex, and scores from the ToL, vocabulary, theory of mind, and shy temperament assessments. The random-effects structure was specified as a random intercept for subject. In the modified version presented here, we ran the model without the effects of diagnosis and the interaction between diagnosis and sex. Before conducting the analysis, we simulated 1000 datasets to determine the power of the study using a random intercept mixed model with a 2-factor within-group variable (condition was a 2-factor within-group variable), with 50 total participants and an intraclass correlation of 0.33. We found approximately 89% power to detect the difference between conditions with an effect size of 0.50, a moderate effect.

We also ran two exploratory models. To anticipate, we had predicted better scores in the audiovisual condition but, instead, obtained better scores in the audio condition. To explore this finding, we split the omnibus score into one for the gestured modality (maximum possible = 11) and one for the spoken modality (maximum possible = 11). We then ran a linear mixed model that included the original model variables plus the additional vocabulary x condition, modality, and modality x condition variables.

In a second exploration, we asked whether overall language ability predicted the expository discourse score. In effect, this is the same question we posed in the registered version of the project but abandoned because we were unable to recruit a sizeable

cohort of children with developmental language disorder before COVID-19. Instead of including the diagnostic category—language disorder or typical language development—in the model, we regressed children's average explanation score (across audio and audiovisual conditions) on their scores on the TNL-2. The TNL-2 was used to group children into DLD or TD categories in the larger longitudinal project. There is evidence that developmental language disorder is a spectrum condition, not a categorical one (Lancaster & Camarata, 2019; Dollaghan, 2004); thus, considering the scores of children who potentially have developmental language disorder on a continuum with those of children who have typical language development is a valid approach.

Finally, we ran a confirmatory model, retesting our primary hypotheses with a linear model.

Results

Descriptive Data

The children's performance on the measures that served as independent variables in the statistical models appears in Table 1. Note from the min-max information that there was a reasonable range of scores on all measures for use in the statistical models. The exact distributions are plotted in the Supplemental Materials (Figures S8 through S18) available at this link OSF | Children's Vocabulary Project; Remote Communication).

Table 1. ***Summary statistics for scores that serve as predictors of expository discourse performance***

Construct	Measure	Score	Mean (sd)	Median	Min-Max
Receptive & Expressive Language	TNL-2	Omnibus Standard Score	104.74 (14.96)	108	72-127
Receptive Vocabulary	NIH PVT	Uncorrected Standard Score	76.76 (6.97)	76.5	60-89
Planning & Problem Solving	ToL	Proportion Extra Moves Score	0.91 (0.32)	0.93	0.14-1.71
Theory of Mind	ToMI-2	Composite Mean	16.95 (1.94)	16.84	12.42-19.89
Shyness	TMCQ	Shyness Total	13.28 (3.41)	13	5-20

Note: TNL-2 = *Test of Narrative Language-2nd edition*, NIH PVT = *NIH Toolbox Picture*

Vocabulary Test, ToL = *Tower of London*, ToMI-2 = *Theory of Mind Inventory-2nd edition*, TMCQ = *Temperament in Middle Childhood Questionnaire*.

Before proceeding, we examined the relationships between the independent variables. The univariate correlations appear in Figure 1. As expected for two language measures, the TNL-2 and NIH Toolbox PVT scores were highly and positively correlated. The TNL-2 scores were also moderately correlated with the ToMI-2 scores. Higher language scores were associated with better theory of mind. Higher language scores were weakly correlated with the shyness scores. Shyer children had lower language scores on the TNL-2 than more outgoing children. The NIH Toolbox PVT scores were moderately correlated with age; higher vocabulary scores were associated with older ages. There were also weak correlations between vocabulary scores and scores on the TMCQ-shy and the ToMI-2. Children with larger vocabularies tended to be less shy and to have a stronger theory of mind.

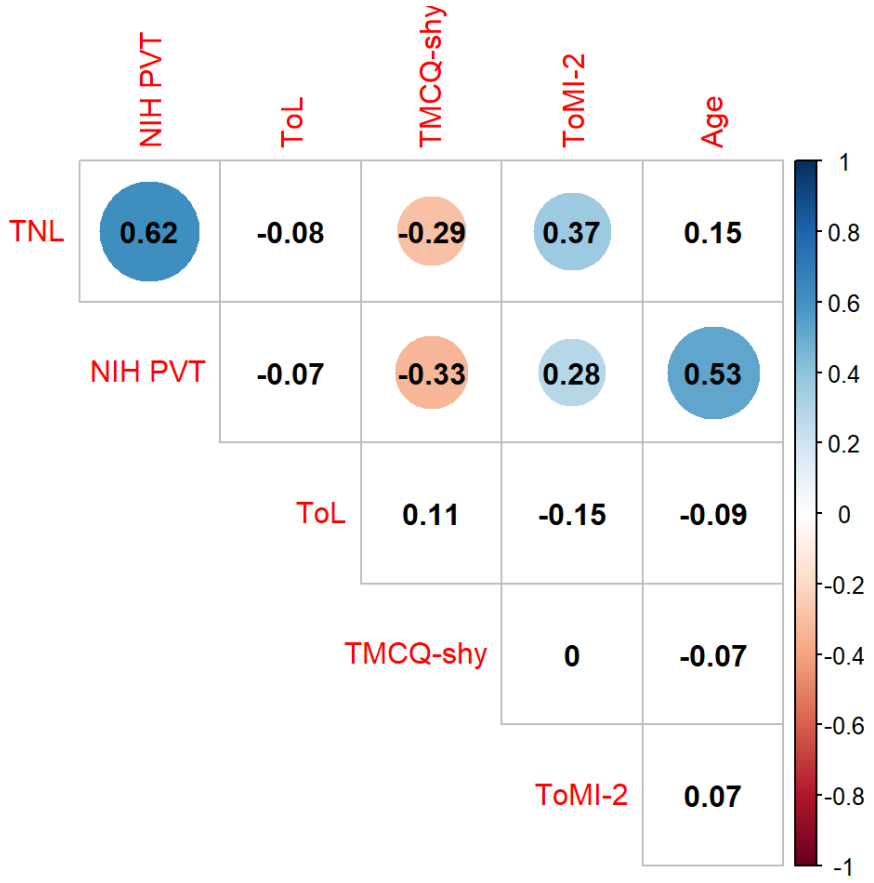

Figure 1. ***Matrix of univariate correlations between Predictor Variables. Cells with circles indicate a significant correlation (p < .05). Figure created via code in Wei and Simko (2017).***

We derived variance inflation estimates (VIF) to determine multicollinearity (Choueiry, 2021; Fox & Weisberg, 2019). The VIF is equal to 1 when a given independent variable is orthogonal to the other independent variables. VIF values between 5 to 10 are considered large and indicative of multicollinearity. To anticipate, we ran models with language measured by the NIH Toolbox PVT or the TNL-2. In either case, VIF estimates were ≤1.775; thus, we could proceed with the models as planned. The detailed results of the VIF analysis appear in the Supplemental Materials.

Discourse

The expository discourse scores are plotted in Figure 2. Scores ranged widely, from as low as 27% of total possible points in the audiovisual condition to as high as 87% in the audio condition.

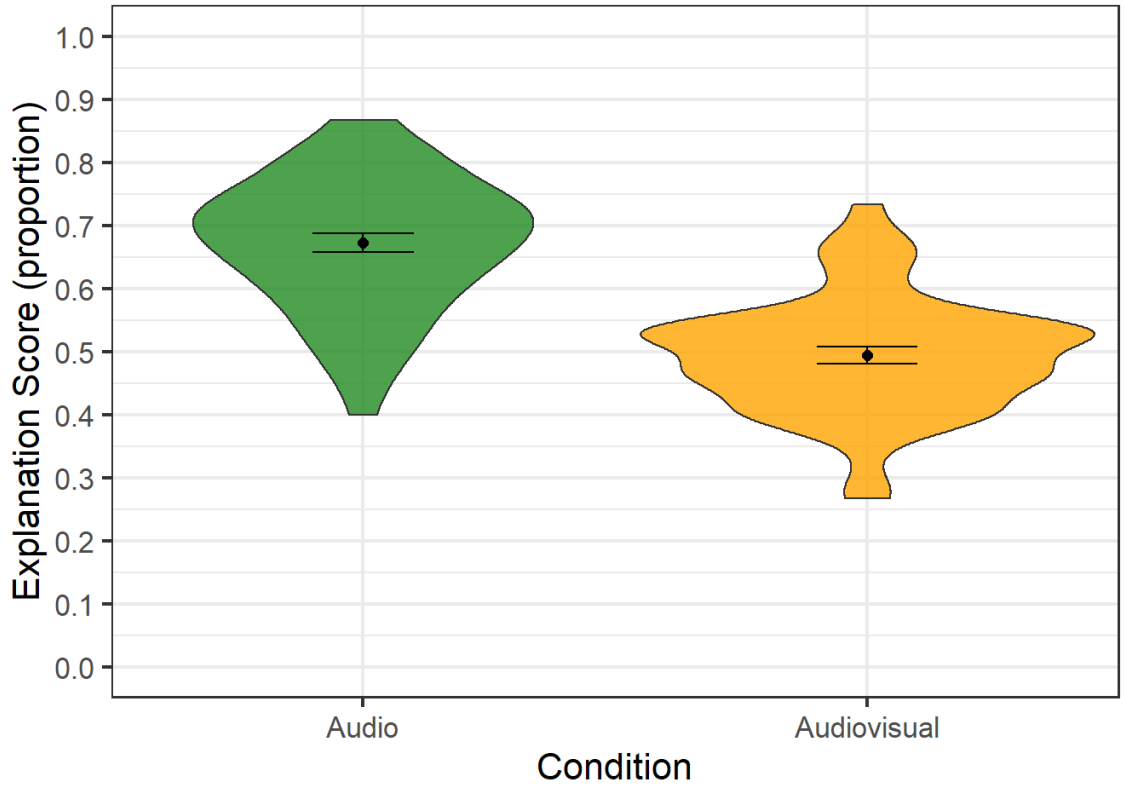

Figure 2. ***Explanation scores (proportion) as a function of Condition (Audio vs. Audiovisual). Diamonds represent the group average and error bars +/- 1 SE. Violins show the distribution of Explanation scores across children.***

The outcome of the model predicting expository discourse performance indexed by

mean explanation scores appears in Table 2. Performance varied with condition; however, the effect was the opposite of our prediction. We found that children's explanation scores were significantly *lower* in the audiovisual (b=0.493) than in the audio (b=0.671) condition. As predicted, there was a significant effect of vocabulary. Children with larger vocabularies had significantly higher explanation scores. For instance, the average explanation score for a child with below-average vocabulary (i.e., 69.79; -1 SD below mean) was 0.55. The average explanation score for a child with an above-average vocabulary (i.e., 83.73; +1 SD above mean) was 0.62.

Table 2. ***Results of Linear Mixed Model Evaluating Predictors of Discourse Performance***

Variable	Estimate	SE	df	t	Pr(>|t|)
Intercept	0.582266	0.011192	43	52.023	< 0.00000001
Condition	-0.177751	0.013307	49	-13.358	< 0.00000001
Sex	-0.009997	0.023173	43	-0.431	0.6683
Age	-0.001828	0.002686	43	-0.680	0.4999
Vocabulary	0.005246	0.002130	43	2.463	0.0179
Tower of London	- 0.013447	0.035539	43	-0.378	0.7070
Theory of Mind	0.010406	0.006326	43	1.645	0.1073
Shyness	-0.002423	0.003554	43	-0.682	0.4990

Note: SE = standard error, df = degrees of freedom, Pr = probability

The remaining predicted effects were not obtained. Among these, it is notable that the explanation scores did not vary with performance on the ToL, the same task the participants were attempting to explain. Thus, any problems in explaining the task were *not* the result of an inability to understand its rules or goals.

To further understanding of the discourse performance, we examined variation by item (Table 3). The children seldom framed their discourse with salutations, openings, or closings. At the other extreme, nearly all children specified spatial and temporal information when conveying how to solve the ToL.

Exploratory analyses

The planned analyses revealed, contrary to prediction, that the children performed better in the audio condition than in the audiovisual condition. Before drawing any conclusions, we examined the data anew. First, we asked whether the condition effect varied with vocabulary knowledge to determine whether the decrease in performance in the video compared to the audio condition was greater for children with better vocabularies. A significant condition x vocabulary interaction could suggest

that the lower explanation scores in the audiovisual condition indicate *more* mature behavior than higher scores(i.e., children with better vocabularies may be more aware of the increased common ground in the audiovisual condition and therefore decrease how much they explain). Next, we asked whether the condition effect varied with modality to determine whether the more robust overall performance in the audio condition was primarily the result of children using more spoken communication when on the phone. The results appear in Table 4.

Table 3: ***The proportion of participants who conveyed each item.***

Item	Audiovisual	Audio
Opening salutation	.22	.10
Closing salutation	.16	.04
Rule 1: boards must look alike	.42	.66
Rule 2: use fewest moves possible	.22	.20
Rule 3: number of beads must not exceed height of peg	.26	.32
Rule 4: move one bead at a time	.14	.30
Problem 1	.20	.54
Problem 2	.22	.60
Problem 3	.24	.50
Problem 4	.20	.46
Problem 5	.18	.38
Specific word for bead	.46	.66
Specific word for peg	.36	.66
Specific word for spatial information	.82	.96
Specific word for sequential information	.78	.84

Table 4: ***Results of Linear Mixed Model Evaluating Predictors of Discourse Performance***

Variable	Estimate	SE	df	t	Pr(>|t|)
Intercept	0.582266	0.011192	43	52.023	<0.0001
Condition	-0.177751	0.013132	48	-13.536	<0.0001
Sex	-0.009997	0.023173	43	-0.431	0.6683
Age	-0.001828	0.002686	43	-1.680	0.4999
Vocabulary	0.005246	0.002130	43	2.463	0.0179
Tower of London	-0.013447	0.035539	43	-0.378	0.7070
Theory of Mind	0.010406	0.006326	43	1.645	0.1073
Shyness	-0.002423	0.003554	43	-0.682	0.4990
Condition x Vocabulary	-0.002895	0.001902	48	-1.522	0.1346

Note: SE = standard error, df = degrees of freedom, Pr = probability

Does the Condition Effect Vary with Vocabulary Knowledge?

The significant effect of Vocabulary, b=0.005, t(43)=2.463, p=0.018 did not vary significantly between conditions, b=-0.003, t(48)=-1.522, p=0.135. In other words, the effect of condition, b=-0.178, t(48)=-13.536, p=0 (i.e., higher explanation scores in the audio than the audiovisual condition) was similar for children with smaller vocabularies and children with larger vocabularies.

Does the Condition Effect Vary with Modality?

We credited the child one point per the first 11 rubric items in the original analysis, whether produced in words, gestures, or both. Here, to discern any differential responses to condition according to the modality of the response, we instead awarded the child one point for each item conveyed with words (11 points maximum) and one point for each item conveyed with gesture(11 points maximum). Children responded primarily by demonstration gestures, scoring on average 5.66 points (SD=1.12) in the audio condition and 6.37 points (SD=1.13) in the audiovisual condition. Children only occasionally responded using deictic gestures, scoring on average 1.35 (SD=1.73) in the audio condition and 1.56 out (SD=2.05) in the audiovisual condition. In addition, children rarely (if ever) responded using representational gestures, scoring on average 0.34 (SD=1.02) in the audio condition and 0.14 (SD=0.4) in the audiovisual condition. Given the relative infrequency of deictic and representational gestures, we did not repeat our analyses separately for each type of gesture. These distributions, however, indicate that children's gesture explanation scores primarily reflect their ability to demonstrate actions on the objects used to complete the task.

Note that, when broken apart by modality, the children's explanation scores were highly correlated, b=0.724, t(48)=2.887, p=0.006. For every one-point increase in children's gestured explanation score, there is a 0.724 increase in their spoken explanation score.

We ran a linear mixed model that included the original model variables plus the modality and modality x condition variables. The results appear in Table 5.

There was a significant effect of condition, b=-0.037, t(97.997)=-2.134, p=0.035. As before, the explanation scores were *lower* in the audiovisual (b=0.4205) than the audio (b=0.4575) condition. Also as before, vocabulary was a significant predictor, b=0.007, t(42.999)=2.747, p=0.009. There was a significant effect of modality. Children's gesture explanation scores were *higher* (b=0.5755) than their spoken explanation scores (b=0.3025). The effect of condition was qualified by a condition x modality interaction, b=0.232, t(97.997)=6.614, p=0. This captures a reversal in the effect of condition.

Children's spoken explanation scores are *lower* in the audiovisual than the audio condition (b=-0.081), while their gestured explanation scores are *higher* in the audiovisual than the audio condition (b=0.151). In both conditions, however, children's gestured explanation scores are *higher* than their spoken explanation scores. This difference is larger in the audiovisual condition (b=0.389) than the audio condition (b=0.157) (Figure 3).

Table 5: ***Results of Linear Mixed Model Evaluating Predictors of Discourse Performance***

Variable	Estimate	SE	df	t	Pr(>|t|)
Intercept	0.439048	0.015628	43.152702	28.094	<0.0001
Condition	-0.037414	0.017528	97.996544	-2.134	0.03530
Sex	-0.002562	0.027175	42.998779	-0.094	0.92531
Age	-0.003715	0.003150	42.998779	-1.179	0.24484
Vocabulary	0.006863	0.002498	42.998779	2.747	0.00874
Tower of London	-0.020329	0.041675	42.998779	-0.488	0.62816
Theory of Mind	0.004530	0.007418	42.998779	0.611	0.54464
Shyness	-0.003506	0.004168	42.998779	-0.841	0.40497
Modality	0.272788	0.024900	48.998357	10.955	<0.0001
Condition x Modality	0.231879	0.035057	97.996544	6.614	<0.0001

Note: SE = standard error, df = degrees of freedom, Pr = probability

Language Ability as a Predictor of Expository Discourse

We already have some evidence that language ability influences expository discourse, given that vocabulary scores were a significant predictor. Next, we determined whether the effect was limited to vocabulary or extended to language ability more broadly defined. Because performance on the NIH PVT and the TNL-2 were significantly correlated, r = .62, we removed vocabulary from the model and replaced it with the TNL-2 scores. The results appear in Table 6.

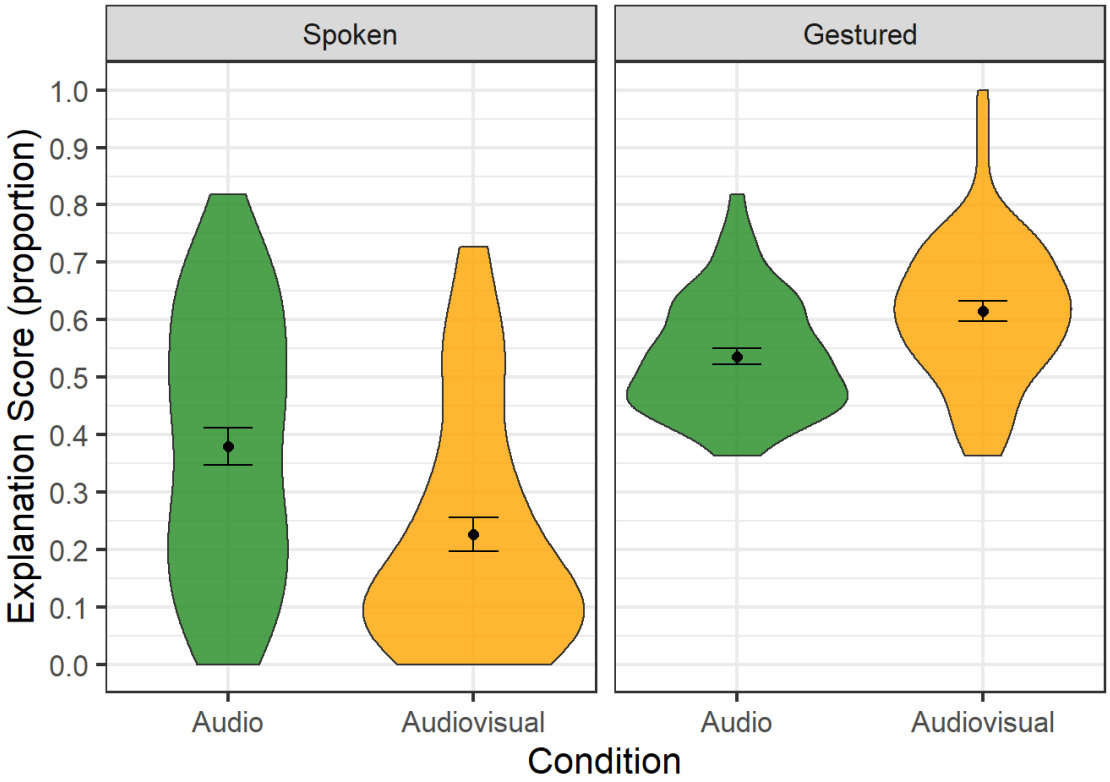

Figure 3: ***Explanation scores (proportion) as a function of Condition (Audio vs. Audiovisual) and modality (Spoken vs. Gestured). Diamonds represent the group average and error bars +/- 1 SE. Violins show the distribution of Explanation scores across children.***

Table 6: ***Results of Linear Mixed Model Evaluating Predictors of Discourse Performance***

Variable	Estimate	SE	df	t	Pr(>|t|)
Intercept	0.5819619	0.0095073	42	61.212	<0.00000002
Condition	-0.1777509	0.0130155	48	-13.657	<0.00000002
Sex	-0.0082890	0.0195446	42	-0.424	0.6737
Age	0.0002123	0.0019940	42	0.106	0.9157
Language	0.0035786	0.0007266	42	4.925	0.0000136
Tower of London	-0.0108946	0.0301276	42	-0.362	0.7195
Theory of Mind	0.0043402	0.0055416	42	0.783	0.4379
Shyness	-0.0009602	0.0029443	42	-0.326	0.7459
Sex x Language	-0.0011283	0.0013768	42	-0.819	0.4171

Note: SE = standard error, df = degrees of freedom, Pr = probability

As expected by now, there was a significant effect of condition. Children's explanation scores were significantly lower in the audiovisual (b=0.493) than in the audio (b=0.671) condition. There was a significant effect of language ability: Children with a higher score on the TNL-2 had significantly higher explanation scores. For instance, the average explanation score for a child with a below-average TNL-2 score (i.e., 89.78; -1 SD below mean) was 0.52. The average explanation score for a child with an above-average TNL score (i.e., 119.7; +1 SD above mean) was 0.64. There were no other significant effects.

Confirmatory Analyses

In our planned analysis, we found a significant effect of condition (audio vs. audiovisual). This analysis, however, controlled for individual differences in our predictor variables (i.e., it was the effect of condition for a child with average performance on each measure). Moreover, the model did not include the full random effects structure (Barr et al., 2013). We could not include a random slope for condition because the model would overfit the data (the number of random effects would match the number of data points per participant - 2). Therefore, we repeated our analyses using the difference between children's explanation scores in each condition and without the predictor variables. This linear model allowed a more direct test of Hypothesis 1 - testing whether explanation scores differ for children in general, and the use of a linear model provided a more transparent way to calculate p-values and effect sizes without the need for random effects. Consistent with the results of our planned analysis, children's total explanation score was significantly lower in the audiovisual condition than the audio condition, b=-0.178, t(49)=-13.358, p=<.001. The condition effect accounted for 78.8% of the variance in children's explanation scores. We fit a second linear model in which we regressed children's explanation scores (averaged across Conditions) on sex, vocabulary, ToL performance, TMCQ-Shyness, ToMI-2, and age. The significant effect of vocabulary was confirmed, b=0.007, t(43)=2.457, p=0.018. Children with larger vocabularies have significantly higher explanation scores. For instance, the average explanation score for a child with below-average vocabulary (i.e., 69.79; -1 SD below mean) was 0.75. The average explanation score for a child with above-average vocabulary (i.e., 83.73; +1 SD above mean) was 0.85. Controlling for the other predictor variables, the size of children's vocabulary accounted for 12.3% of the variance in children's explanation scores. No other predictors were significant, p's > 0.112. The details appear in the Supplemental Materials.

Discussion

Remote communicative contexts are part of everyday social, familial, and academic interactions for the modern child. However, much of what we know about the child's ability to meet the informational needs of a communicative partner is based on data

collected during face-to-face interactions. This study adds to the sparse extant literature on communicative success in remote contexts.

Thanks to the rise in video-chat apps available in home and school settings, remote communication increasingly involves a shared visual context. Traditional phone communication does not. Thus, in this study, we were interested in the completeness of children's remote discourse, its variation with the presence or absence of visual context, the extent to which children modified their use of gestural communication when moving between contexts, and the characteristics that predict children's communicative success. Below we organize our findings into two main categories: how remote communication varies with context and how it varies with child characteristics.

Variation Associated with Context Demands

Remote communication places a high demand on speakers. They must infer and then meet their listeners' need for information without many of the cues available in face-to-face communication. Given that the audiovisual condition reinstates some of these cues, we predicted that the children's overall explanation scores would be higher in the audiovisual condition than in the audio condition. Specifically, we anticipated that the children would use gestures to supplement their spoken messages. Relative to the audio condition, the children did gesture more information in the audiovisual condition, but, at the same time, they provided less information in the spoken modality; thus, the overall rubric score was higher in the audio condition, contrary to prediction. This finding is consistent with Cameron and Lee (1997), who reported that 3-to-8-year-olds provided more detail and specificity while speaking on the phone than in person. The children responded to their listeners' needs by adjusting to the listener's need for spoken input.

The particular items that the children most often included in their discourse also illustrate their sensitivity to listeners' needs. Consider, for example, the low rate of salutations and the high rate of spatial and temporal content. The pragmatic framing of the discourse with salutations was uncommon, but such niceties are not necessary for explaining the task (and perhaps awkward given that the partner was not present). In contrast, when and where to move the beads were details that nearly every child included, and this information was essential for solving the task.

Thus, just as they do when face-to-face (e.g., Akhtar et al., 1996; Mori, & Cigala, 2016; Nadig & Sedivy, 2002; Nilsen, & Fecica, 2011; Shatz & Gelman, 1973), children demonstrate adaptations in content and modality according to their listener's needs when communicating remotely. When the remote context lacked shared visual reference, they enhanced the clarity and completeness of their spoken messages. When their partner was able to see them, they offloaded some of their verbal explanation into gestures. In her nuanced description of discourse in a year-five classroom, Taylor

(2014, p. 416) wrote:

"It is important to emphasize that modes other than language are not simply additional contextual information but part of an enmeshed nexus of many modes used in conjunction with one another for the purpose of making meaning. All modes are potentially available for making meaning, within the constraints of our social world. The mode selected by the communicator is the one judged by them to be the most apt and expedient at that moment in time."

Gesture was, of course, an apt and expedient means of communication in the audiovisual condition. The types of gestures the children used were well suited to the partners' needs. All of the participants used demonstration gestures. These were hand gestures that resulted from manipulating the materials, in other words, moving the beads. As the primary goal was to teach the partner how to move the beads, demonstration gestures were an effective means of explanation.

Of course, the demonstration gestures were apt and expedient for the partner in the audiovisual condition only. Nonetheless, the children gestured more items than they presented in words even in the audio condition. We do not take this as a counter to the conclusion that they were sensitive to the partners' needs. Instead, gestures can be apt and expedient for the speaker as well as the listener (Goldin-Meadow, 2003). Speakers use gestures with exceptionally high frequency when communicating spatial information (Alibali, 2005). When explaining a visual-spatial task, working out the problem by moving the hands through space is an excellent strategy for thinking through the steps one must convey. The children likely gestured in the audio condition (and to some extent in the audiovisual condition) because the gestures helped them explain the task. Had the children been given repeated practice with the ToL, we would predict less reliance upon gestures that involved demonstrations on objects and more free-handed gestures. As it were, their high use of demonstration gestures was consistent with their status as novice ToL solvers (Roth, 2002). We turn now to other characteristics that were related to the success of their remote discourse.

Variation Associated with Language Ability

Motivated by previous work on the influence of language, theory of mind, and temperament on communicative success, we tested the predictive utility of language scores, theory of mind ratings, and temperament—specifically shyness—ratings in our models of remote discourse success. However, we first measured the inter-dependence of these predictors. Given the equivocal reports of relationships between shyness and theory of mind, some reporting a negative relationship (Banerjee & Henderson, 2001; DeRosnay et al., 2014; Walker, 2005) and others a positive relationship (Mink et al., 2014; Wellman et al., 2011), it is noteworthy that we found neither in the current sample.

That said, we did find a relationship between shyness and language; specifically, the shyer children in our sample tended to have lower vocabulary and receptive/expressive language scores than more outgoing children. Spere et al. (2004) compared the receptive vocabulary scores of four-year-olds grouped as shy or not shy and found the shyer children to have significantly lower scores. Here we extend these findings to an older cohort. Some have speculated that their reticence to speak masks the language competency of shy children (see summary in Coplan & Evans, 2009). Although this could be the case, we found a negative relationship between shyness and performance on a receptive vocabulary test (which does not require spoken responses), a finding at odds with the masked language competency hypothesis. Another possibility is that the lower test scores on both the TNL-2 and the NIH-PVT reflect more test anxiety on the part of the shyer children, but we think this is unlikely given that previous work has established the validity of standardized tests administered to shy children. Specifically, shy children did not perform better on language tests administered in the home by a familiar adult than on those same tests administered at school by an unfamiliar adult (Spere et al., 2009).As has been previously proposed (Spere et al., 2004), we think it likely that children who are shy limit their opportunities for language learning by refraining from social-communicative interactions. Shyness (or temperament more broadly measured) may be a source of individual differences in children's language outcomes.

Language, as measured by the TNL-2 and, to a lesser extent, by the NIH-PVT, was also correlated with theory of mind. Milligan et al. (2007) also reported a positive relationship between language and theory of mind with an overall effect size of .43 among children below seven. Here we extend that finding to children who are seven to nine and report a similar effect size of .37 (on the TNL-2). The relation between language and theory of mind is likely bidirectional. Children who participate frequently and competently in communicative exchange access multiple opportunities for learning about others' mental states, and conversely, children who are skilled at mind-reading may learn mental state vocabulary and hone their social language skills upon realizing that their listener is confused, skeptical, interested or bored (De Rosnay et al., 2014). Language ability in the form of complex sentence construction may also aid thought about others' mental states, especially at the relatively older ages tested here, years during which children may be progressing from first order (I suspect he is hungry) to second-order observations (He knows that I suspect he is hungry) (de Villiers, 2007).

Our analyses allowed us to examine the potential effects of language, theory of mind, and shy temperament on discourse, each after controlling for the others. We also examined the effects of sex, ability to solve the ToL, and age. Whether measured as receptive vocabulary or a receptive and expressive narrative ability, language was the only predictor. Children with stronger language abilities produced more complete explanations of the ToL during the discourse task, as evident by their rubric scores.

There were no significant interactions between language and condition or between language and modality. In other words, regardless of their vocabulary knowledge, children tended to offload more information onto gestures when the visual context allowed. The high positive correlation between gesture scores and spoken language scores further supports the conclusion that children's gesture use was a sign of the integrity of their overall communicative competence rather than a way of compensating for communicative weaknesses. These findings are consistent with age-related differences in communicative competence. Older children not only use more complex spoken language but also more complex gestures than younger children (Alamillo et al., 2013). Like teachers who package relevant information into their gestures during classroom lectures (Alibali et al., 2013; Ovendale et al., 2018), children who frequently gesture when sharing a visual context with their interlocutor likely maximize the effectiveness of their message.

We do not dismiss the potential influence of theory of mind or shy temperament in other discourse contexts. Recall that, in the discourse task used here, we told the children that their communication partner did not know how to solve the ToL; thus, we likely reduced the need for mind reading. Moreover, the children did not interact with their partner but, instead, were recorded for later listening or viewing. This situation may have lessened the burden that shyer children may have felt had they been part of an actual exchange. The decision to simulate phone and video chat rather than engage the children in these actual contexts was purposeful. We wanted to control the amount of feedback a listener would provide, but we could not imagine how to do so in a pragmatically appropriate way. By recording the children's discourse, we got around this problem. That said, this strength is also a limitation of the work. The child was at a remove from an actual communicative exchange. Moreover, the child did not receive the scaffolding that the verbal and gestural responses of a listener would have provided, which surely made the discourse task more difficult than usual, perhaps especially so for those with weaker language abilities. Observations of naturalistic remote discourse would be a valuable complement to the work reported here. Also, a comparison between the two types of remote discourse studied here and actual face-to-face discourse would be helpful if we are to understand fully the challenges involved in remote communication.

Finally, we turn to the implications of the language as a predictor of remote discourse skills. This remote discourse task was difficult. None of the children provided 100% of the information we deemed essential. That said, some of the children had particularly poor performance. On average, those whose receptive/expressive language scores fell one standard deviation below the mean provided only 30% of the essential information. Real-world remote communication is likely to be challenging for these children unless their partner provides ample scaffolding in the form of feedback and questions. Given the ubiquity of remote communication in children's lives, it is essential to document how the estimated 9% of children with language disorders (Norbury

et al., 2016) fare in remote contexts and determine the supports needed to ensure adequate remote discourse function.

Conclusions

Among the limitations of this study were our inability to complete it as registered and the lack of ecological validity inherent in a partnerless simulation. That said, we provided evidence of positive relationships between language and theory of mind and negative relationships between language and shyness, extending the extant literature to older children. We also confirmed that, as a group, seven-to-nine-year-olds adjust their discourse to the needs of their remote communication partners. They include essential semantic information, and when they know that their partner does not share their visual context, they still gesture frequently, but they increase their reliance upon the spoken modality. Perhaps the primary contribution of this work is the finding that remote discourse is challenging, even for children as old as nine, and especially so for children who have below-average receptive and expressive language abilities. This finding has important practical implications given that children's communication partners—their friends, families, teachers, and health care providers—are frequently remote. Children with low language abilities may experience functional limitations during remote communication, a context that is increasingly necessary in today's world.

References

Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Development, 67(2),* 635-645. https://doi.org/10.1111/j.1467-8624.1996.tb01756.x

Alamillo, A.R., Colletta, J., & Guidetti, M. (2013). Gesture and language in narratives and explanations: the effects of age and communicative activity on late multimodal discourse development. *Journal of Child Language, 40*, 511-538. https://doi.org/10.1017/S0305000912000062

Alibali, M. W. (2005). Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation, 5*(4), 307-331. https://doi.org/10.1207/s15427633scc0504_2

Alibali, M. W., & Goldin-Meadow, S. (1993). Modeling learning using evidence from speech and gesture. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 203-208). Hillsdale, NJ: Erlbaum.

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes, 15*(6), 593-613. https://doi.org/10.1080/016909600750040571

Alibali, M. W., Young, A.G., Crooks, N.M., Yeo, A., Wolfgram, M.S., … Knuth, E.J. (2013). Students learn more when their teacher has learned to gesture effectively. *Gesture, 13* (2), 210–233. https://doi.org/10.1075/gest.13.2.05ali

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences*, *4*(7), 267-278. https://doi.org/10.1016/S1364-6613(00)01501-1

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology, 35*(5), 1311. https://doi.org/10.1037/0012-1649.35.5.1311

Banerjee, R. & Henderson, L. (2001). Social-cognitive factors in childhood social anxiety. *Social Development,* 10(4), 558-572. https://doi.org/10.1111/1467-9507.00180

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68(3),* 255-278. https://doi.org/10.1016/j.jml.2012.11.001

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General*, *136*(4), 539. doi:10.1037/0096-3445.136.4.539

Buss, A. H., & Plomin, R. (1984). *Theory and measurement of EAS, temperament: early developing personality traits,* pp. 84-104. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Cameron, C. A., & Lee, K. (1997). The development of children's telephone communication. *Journal of Applied Developmental Psychology, 18*(1), 55-70. https://doi.org/10.1016/S0193-3973(97)90014-9

Capone, N. C., & McGregor, K. K. (2004). Gesture development. *Journal of Speech, Language, and Hearing Research, 47*, 173-186. https://doi.org/10.1044/1092-4388(2004/015)

Castelli, F., Frith, C., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain, 125*(8), 1839-1849. https://doi.org/10.1093/brain/awf189

CCSS (Common Core State Standards). (2015). English Language Arts Standards » Speaking & Listening » Grade 4 » 6 | Common Core State Standards Initiative (corestandards.org). Downloaded 1/10/2021.

Choueiry, G. (2021). Correlation vs Collinearity vs Multicollinearity. Correlation vs Collinearity vs Multicollinearity – Quantifying Health Downloaded 3/11/21.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23*(1), 43-71. https://doi.org/10.1016/0010-0277(86)90053-3

Coplan, R. J., & Evans, M. A. (2009). At a loss for words? Introduction to the special issue on shyness and language in childhood. *Infant and Child Development, 18*, 211–215. https://DOI: 10.1002/icd.620

Coplan, R. J., & Weeks, M. (2009). Shy and soft-spoken: shyness, pragmatic language, and socio-emotional adjustment in early childhood. *Infant and Child Development: An International Journal of Research and Practice, 18*(3), 238-254. https://doi.org/10.1002/icd.622

Culbertson, W.C., & Zillmer, E.A. (2005). *Tower of London Drexel University, 2nd Edition,* North Tonawanda, NY: Multi-Health Systems

de Rosnay, M.C., Fink, E., Begeer, S., Slaughter V., & Peterson C. (2014). Talking theory of mind talk: young school-aged children's everyday conversation and understanding of mind and emotion. *Journal of Child Language*, 41(5), 1179-1193. https://doi.org/10.1017/S0305000913000433

de Villiers J. (2007). The Interface of Language and Theory of Mind. *Lingua. International review of general linguistics. Revue internationale de linguistique generale, 117*(11), 1858–1878. https://doi.org/10.1016/j.lingua.2006.11.006

Dollaghan, C. A. (2004). Taxometric analyses of specific language impairment in 3- and 4-year-old children. *Journal of Speech, Language & Hearing Research, 47*(2). https://doi.org/10.1044/1092-4388(2004/037)

Dorval, B., Eckerman, C. O. & Ervin-Tripp, S. (1984). Developmental trends in the quality of conversation achieved by small groups of acquainted peers. *Monographs of the Society for Research in Child Development,* 49(2), Serial No. 206. https://doi.org/10.2307/1165872

Ehrlich, S. B., Levine, S. C., & Goldin-Meadow, S. (2006). The importance of gesture in children's spatial reasoning. *Developmental Psychology*, *42*(6), 1259. https://doi.org/10.1037/0012-1649.42.6.1259

Erstad O, Flewitt R, Kümmerling-Meibauer B, et al. (eds) (2020) *The Routledge Handbook of Digital Literacies in Early Childhood*. Abingdon: Routledge.

Farrar, M. J., & Maag, L. (2002). Early language development and the emergence of atheory of mind. *First Language, 22*(2), 197-213. https://doi.org/10.1177/014272370202206504

Fox, J., & Weisberg, S (2019). *An R Companion to Applied Regression*, Third edition. Sage, Thousand Oaks CA. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Garber, P., & Goldin-Meadow, S. (2002). Gesture offers insight into problem-solving in adults and children. *Cognitive Science, 26*(6), 817-831. https://doi.org/10.1207/s15516709cog2606_5

Gershon, R. C., Gleason, J. B., Michnick Golinkoff, R., Jager Adams, M., Schnipke, D., Hirsh-Pasek, K., Slotkin, J., Manly, J. J., Blitz, D., Beaumont, J. L., Wallner-Allen, K., Weintraub, S. (2013). NIH Toolbox - Cognition - Early Childhood Picture Vocabulary Test (NIH TPVT - Early Childhood) National Institute of Health and Northwestern University. NIH TPVT - Early Childhood - NIH Toolbox - Cognition - Early Childhood Picture Vocabulary Test (mapi-trust.org)

Gillam, R. B., & Pearson, N. A. (2017). *TNL-2: Test of Narrative Language*. Pro-ed.

Gillen, J. (2002). Moves in the territory of literacy? The telephone discourse of three- and four-year-olds. *Journal of Early Childhood Literacy, 2*(1), 21-43. https://doi.org/10.1177/14687984020021002

Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2010). How do preschoolers express cause in gesture and speech?. *Cognitive Development, 25*(1), 56-68. https://doi.org/10.1016/j.cogdev.2009.11.001

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Belknap Press of Harvard University Press.

Henderson, H. A., & Wachs, T. D. (2007). Temperament theory and the study of cognition–emotion interactions across development. *Developmental Review, 27*(3), 396-427. https://doi.org/10.1016/j.dr.2007.06.004

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin, 137*(2), 297. https://doi.org/10.1037/a0022128

Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development, 13*(4), 590-619. https://doi.org/10.1111/j.1467-9507.2004.00285.x

Hutchins, T. and Prelock, P., 2016. Technical manual for the theory of mind inventory-2. Unpublished Copyrighted Manuscript. Available at: theoryofmindinventory.com.

Koeze, E., & Popper, N. (April 7, 2020). The virus changed the way we internet. The New York Times. Downloaded on 1/10/2021 from https://www.nytimes.com/interactive/2020/04/07/technology/coronavirus-internet-use.html.

Lancaster, H. S., & Camarata, S. (2019). Reconceptualizing developmental language disorder as a spectrum disorder: Issues and evidence. *International Journal of Language & Communication Disorders*, *54*(1), 79-94. https://doi.org/10.1111/1460-6984.12433

Larson, C., Gangopadhyay, I., Kaushanskaya, M., & Weismer, S. E. (2019). The relationship between language and planning in children with language impairment. *Journal of Speech, Language, and Hearing Research*, 62(8), 2772-2784. https://doi.org/10.1044/2019_JSLHR-L-18-0367

Lidstone, J. S., Meins, E., & Fernyhough, C. (2012). Verbal mediation of cognition in children with specific language impairment. *Development and Psychopathology.*, *24*(2), 651-660. http://dx.doi.org/10.1017/S0954579412000223

Lundine, J. P., & McCauley, R. J. (2016). A tutorial on expository discourse: Structure, development, and disorders in children and adolescents. *American Journal of Speech-Language Pathology*, 25(3), 306-320. https://doi.org/10.1044/2016_AJSLP-14-0130

Marton, K. (2008). Visuo-spatial processing and executive functions in children with specific language impairment. *International Journal of Language & Communication Disorders*, 43(2), 181-200. https://www.tandfonline.com/doi/citedby/10.1080/13682820701340719?scroll=top&needAccess=true

McClure, E. R., Chentsova-Dutton, Y. E., Barr, R. F., Holochwost, S., & Parrott, W. G. (2015). "Facetime doesn't count": Video chat as an exception to media restrictions for infants and toddlers. *International Journal of Child-Computer Interaction*, 6, 1–6. https://doi.org/10.1016/j.ijcci.2016.02.002

McClure, E. R., Chentsova-Dutton, Y. E., Holochwost, S. J., Parrott, W. G., & Barr, R.

(2018). Look at that! Video chat and joint visual attention development among babies and toddlers. *Child Development, 89*(1), 27-36. https://doi.org/10.1111/cdev.12833

McGregor, K. K., Eden, N., Arbisi-Kelm, T., Foody, M., & Oleson, J. (2019, September 3). Children's Vocabulary Project; Remote Communication. Retrieved from osf.io/cuhy 10.17605/OSF.IO/CUHYX

McGregor, K. K., Eden, N., Arbisi-Kelm, T., Oleson, J., & Pomper, R. (2020, November 6). Children's Vocabulary Project; Remote Communication. Retrieved from osf.io/we8am

Miller, C. A. (2006). Developmental relationships between language and theory of mind. *American Journal of Speech-Language Pathology, 15(2), 142-154.* https://doi.org/10.1044/1058-0360(2006/014)

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78*(2), 622-646. https://doi.org/10.1111/j.1467-8624.2007.01018.x

Mink, D., Henning, A., & Aschersleben, G. (2014). Infant shy temperament predicts preschoolers theory of mind. *Infant Behavior and Development, 37*(1), 66-75. https://doi.org/10.1016/j.infbeh.2013.12.001

Mori, A., & Cigala, A. (2016). Perspective taking: Training procedures in developmentally typical preschoolers. Different intervention methods and their effectiveness. *Educational Psychology Review, 28*(2), 267-294. https://doi.org/10.1007/s10648-015-9306-6

MPFS: Medienpädagogischen Forschungsverbundes Südwest (Media Educational ResearchAssociationnSouthwest) (2018). Kindheit, Internet, Medien Studie (Childhood, Internet, Media Study). Downloaded 3/11/2021 from KIM-Studie 2018 - Kindheit, Internet, Medien - Initiativbüro Gutes Aufwachsen mit Medien (gutes-aufwachsen-mit-medien.de)

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. https://doi.org/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. https://doi.org/10.1016/j.dr.2011.07.001

Neppl, T. K., Donnellan, M. B., Scaramella, L. V., Widaman, K. F., Spilman, S. K., Ontai, L. L., & Conger, R. D. (2010). Differential stability of temperament and personality from toddlerhood to middle childhood. *Journal of Research in Personality, 44*(3), 386-396. https://doi.org/10.1016/j.jrp.2010.04.004

Nippold, M. A., Mansfield, T. C., Billow, J. L., & Tomblin, J. B. (2008). Expository discourse in adolescents with language impairments: Examining syntactic development. *American Journal of Speech-Language Pathology*, 17(4), 356-366. https://doi.org/10.1044/1058-0360(2008/07-0049)

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., ... & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. *Journal of Child Psychology and Psychiatry, 57(11),* 1247-1257. https://doi.org/10.1111/jcpp.12573

Ovendale, A., Brookes, H., Colletta, J. M., & Davis, Z. (2018). The role of gestural polysigns and gestural sequences in teaching mathematical concepts: The case of halving. *Gesture, 17*(1), 128-157. https://doi.org/10.1075/gest.00013.ove

Pine, K. J., Lufkin, N., & Messer, D. (2004). More gestures than answers: children learning about balance. *Developmental Psychology, 40*(6), 1059. https://doi.org/10.1037/0012-1649.40.6.1059

Pinto, G., Tarchi, C., Gamannossi, B. A., & Bigozzi, L. (2016). Mental state talk in children's face-to-face and telephone narratives. *Journal of Applied Developmental Psychology, 44,* 21-27. https://doi.org/10.1016/j.appdev.2016.02.004

Prior, M., Smart, D., Sanson, A. N. N., & Oberklaid, F. (2000). Does shy-inhibited temperament in childhood lead to anxiety problems in adolescence?. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(4), 461-468. https://doi.org/10.1097/00004583-200004000-00015

Research Registry 3425 (2017). The Dynamics of Word Learning, McGregor, P.I. Browse the Registry - Research Registry

Roello, M., Ferretti, M. L., Colonnello, V., & Levi, G. (2015). When words lead to solutions: Executive function deficits in preschool children with specific language impairment. *Research in Developmental Disabilities,* 37, 216-222. https://doi.org/10.1016/j.ridd.2014.11.017

Roth, W. M. (2002). From action to discourse: The bridging function of gestures. *Cognitive Systems Research, 3*(3), 535-554. https://doi.org/10.1016/S1389-0417(02)00056-6

Schmidt, L. A., & Tasker, S. L. (2000). Childhood shyness: Determinants, development and 'depathology'. In W.R. Crozier, (Ed.) *Shyness: Development, consolidation and change,* 30-46. Routledge.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London,* 298(1089), 199-209. https://doi.org/10.1098/rstb.1982.0082

Shatz, M, & Gelman, R. (1973). The development of communication skills:Modifications in the speech of children as a function of listener. *Monographs of the Society for Research in Child Development,* 38(5), Serial No. 152. https://doi.org/10.2307/1165783

Simonds, J., & Rothbart, M. K. (2006). Temperament in Middle Childhood Questionnaire. *Downloaded in http://www. bowdoin. edu/~ sputnam/rothbart-temperament-questionnaires*.

Smith Watts, A. K., Patel, D., Corley, R. P., Friedman, N. P., Hewitt, J. K., Robinson, J. L., & Rhee, S. H. (2014). Testing alternative hypotheses regarding the association between behavioral inhibition and language development in toddlerhood. *Child Development, 85*(4), 1569-1585. https://doi.org/10.1111/cdev.12219

Spere, K. A., Evans, M. A., Hendry, C. A., & Mansell, J. (2009). Language skills in shy and non-shy preschoolers and the effects of assessment context. *Journal of Child Language, 36*(1), 53. https://doi:10.1017/S0305000908008842

Spere, K. A., Schmidt, L. A., Theall-Honey, L. A., & Martin-Chang, S. (2004). Expressive and receptive language skills of temperamentally shy preschoolers. *Infant and Child Development: An International Journal of Research and Practice, 13(2),* 123-133. https://doi.org/10.1002/icd.345

Taylor, R. (2014). Meaning between, in and around words, gestures and postures–multimodal meaning-making in children's classroom discourse. *Language and Education, 28*(5), 401-420. https://doi.org/10.1080/09500782.2014.885038

Uccelli, P., Demir-Lira, Ö. E., Rowe, M. L., Levine, S., & Goldin-Meadow, S. (2019). Children's early decontextualized talk predicts academic language proficiency in mid adolescence. *Child Development, 90*(5), 1650-1663. https://doi.org/10.1111/cdev.13034

Veinott, E. S., Olson, J., Olson, G. M., & Fu, X. (1999, May). Video helps remote work: Speakers who need to negotiate common ground benefit from seeing each other. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 302-309). https://doi.org/10.1145/302979.303067

Walker, S. (2005). Gender differences in the relation between children's peer-rated social competence and individual differences in theory of mind. *Journal of Genetic Psychology,* 166(3), 297-312. https://doi.org/10.3200/GNTP.166.3.297-312

Wechsler, D. (1999). Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: The Psychological Corporation.

Wei, T. & Simko, V. (2017). R package "corrplot": Visualization of a Correlation Matrix (Version 0.84). Available from https://github.com/taiyun/corrplot

Wellman, H. M., Lane, J. D., LaBounty, J., & Olson, S. L. (2011). Observant, nonaggressive temperament predicts theory-of-mind development. *Developmental Science, 14*(2), 319-326. https://doi.org/10.1111/j.1467-7687.2010.00977.x

Zhao, L., Wang, J. J., & Apperly, I. A. (2018). The cognitive demands of remembering a speaker's perspective and managing common ground size modulate 8-and 10-year-olds' perspective-taking abilities. *Journal of Experimental Child Psychology, 174,* 130-149. https://doi.org/10.1016/j.jecp.2018.05.013

Data, Code and Materials Availability Statement

The raw data, analysis code, plots of data distributions, videotaped examples, and a link to the registration appear in McGregor et al. (2020, OSF | Children's Vocabulary Project; Remote Communication)

Ethics Approvals and Consent

Ethics approval was obtained from the ethics committee of the Boys Town National Research Hospital. All participants and their parents gave informed written consent before taking part in the study.

Authorship and Contributorship Statement

KKM secured funding for the study, conceived of the study, designed the study, supervised data collection and analysis, wrote the introduction, results, and discussion section and revised the methods section. RP contributed to the planning of the statistical analysis, conducted the statistical analysis and prepared the figures. NE collected data, coded the discourse data, and wrote the methods section as it pertained to the discourse data. TA-K collected data and wrote the participants section of the Methods. NO collected data and wrote the Tower of London description in the Methods. SG conducted reliability checks on the discourse data and assisted with the literature search. ES assisted with statistical analysis and wrote the statistical analysis section of the Methods. All authors approved the final version of the manuscript and agree to be

accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgements

We thank the children and their families for participation and Jacob Oleson for consultation on the statistical analyses. We give additional thanks to the family who agreed to share the video-taped examples of the discourse task. This work was supported by the National Institutes of Health-NIDCD, grant number 2R01DC011742-06. The funder had no part in the study design; the collection, analysis, and interpretation of data; the writing of the report; or the decision to submit the article for publication.

Appendix

Discourse Instructions for Participants

"Now that you know how to play this game, I'm going to ask you to teach two of my friends."

Audio

"This is X (*photo of woman A or B on phone*). She couldn't be here today but she told me that we could record the directions. Then she will borrow the game later and try it out. She doesn't have a computer screen so she won't be able to see you but she DOES have her cell phone, so she can listen to your recording. I'll start the recording (*make a big fuss about the microphone*). You explain to her how to play the game, and she will listen to what you say. Be very careful to tell her what the game looks like, what the rules are, and exactly how to play."

Now that they know about the game, you can help them solve five of the problems. I'll get the game set up for you each time. "You are the teacher. Tell X how to make this one (*pt to child's board*) look like this one (*pt to target board further away*)."

Repeat for items p, p, 1, 2, 3.

"That was great! Okay, I'm going to ask you to do that one more time for a different friend."

Audiovisual

"This is X (*photo of woman A or B facing screen*). She couldn't be here today but she told me that we could video the directions. Then she will borrow the game later and try it out. She has a computer so she WILL be able to hear you AND see you. I'll start

the recording (*make a big fuss about the video camera*). You explain to her how to play the game. She will watch and listen to what you say. Be very careful to tell her what the rules are, and exactly how to play."

Now that they know about the game, you can help them solve five of the problems. I'll get the game set up for you each time. "You are the teacher. Tell X how to make this one (*pt to child's board*) look like this one (*pt to target board further away*)."

Repeat for items p, p, 1, 2, 3.

Debriefing

You did so well being the teacher. Sometimes it is harder to be the teacher when the person you are talking to can't see you. My friends are going to study your recordings to see how children talk to people who can't see them, like when you talk on a phone. They won't really be watching to learn the game, that part was just pretend.

Discourse Scoring

One point was awarded for each item that the child completed successfully. Except for the final four items, the child was credited for successful completion via words OR gestures. For the five trials (items P1, P2, 1, 2, 3 on the ToL), we assumed that the listener had the visual context of the two boards, one in start position and one in target position. Total scores could range from 0 – 15.

Greets the listener (2 points):

- Hello
- Goodbye

Verbal credit was given if the child offered the listener a stated "hello" or "goodbye." Iconic gestures were credited if the child waved to greet or bid farewell to the listener.

Rules (4 points)*:*

- Explains that the two boards must be alike
- Explains that you must use as few moves as possible
- Explains that no peg can have more beads than it can hold
- Explains that you can move only one bead at a time

Children earned one point per rule if they provided an adequate explanation with words and/or gestures.

Examples for *both boards must look alike*:

- "You have to copy the teacher's board."
- "Try to do the same pattern."
- The child set-up the teacher's board with an example pattern, and then showed the listener how to manipulate the beads in order to make the other board match.

Examples for *complete the pattern in as few moves as possible*:

- "Solve in the least amount of moves"
- "Use less moves as possible"
- "Whoever has less of the movings wins."
- The child used a hand gesture to illustrate moving the beads, while using the word 'moves,' and holding up two fingers to show that the example problem required only two moves.

Examples for *no peg can contain more beads than it can hold:*

- "The small one can hold one *bead,* the middle one can hold two, and the tall one can hold three,"
- "The large *peg* can hold three, the middle peg can hold two and the small one can hold one." The child manipulated the beads on the board to show the listener the maximum bead amounts for each peg.
- The child pointed to a peg as they stated how many beads could be placed on it.

Examples for *move only one bead at a time.*

- "You can only move one bead at a time."
- "You can only take one off the peg at a time."
- Children often explained this rule via demonstration gestures by showing the various ways in which this rule can be broken, in the same way it was presented to them prior to completing the standardized portion of the test (i.e., lifting two beads off the pegs with one hand/lifting two beads off the pegs with two hands/lifting one bead off the peg and placing it on the table, and then proceeding to take another bead off a peg).

Example Problems (5 points):

- Gives enough information for listener to correctly solve trial P1, minimum number of moves is not required
- Gives enough information for listener to correctly solve trial P2, minimum

number of moves is not required
- Gives enough information for listener to correctly solve trial 1, minimum number of moves is not required
- Gives enough information for listener to correctly solve trial 2, minimum number of moves is not required
- Gives enough information for listener to correctly solve trial 3, minimum number of moves is not required

To earn verbal points in the absence of gesture, the child was required to provide the listener with sufficient information to solve the problem accurately using only spoken language. Credited instructions included referents and/or descriptive words for each move within the problem. Acceptable descriptors for identifying the target peg included distinctions in size (e.g., short/tall/little/long/middle) and body-oriented directional terms (e.g., left/middle/right.) Credit was also given if the child identified where a bead should be placed by stating the color of the bead already atop the target peg (e.g., Put the green one on the red one), or by stating the color of the bead that was in the target position prior to the previous move (e.g., "Put the red one where the blue one used to be"). In the absence of gesture, environmental-oriented directional terms (e.g., front/back, first/last) were not considered specific enough to describe target pegs (e.g., "Put the blue bead on the front peg and move the red bead to the back one" was too vague to receive credit because it is not clear which is front and which is back).

To earn credit via demonstration gestures, the child needed to move the beads from the starting position to the target position without breaking any rules. Credit was given for the successful completion of the problem, despite move count.

Children received credit if they combined verbal description and gesture to clearly convey content:

- "The red one goes on this peg" (pointing to peg)
- "See this blue bead [participant moves bead closer to the listener], it goes on this peg".
- "I move the red one to the small peg, and that's one move [holds up one finger]", and "I'm going to switch these around [moves hand from left to right to indicate switching]."

Specific content words (4 points):

- Refers to the beads with a relevant word (e.g., bead, ball) at least once
- Refers to pegs with a relevant word (e.g., peg, stick, stand) at least once

- Uses a relevant directional term (e.g., here/there, right/left) at least once
- Uses a relevant sequential term (e.g., first, next, then) at least once

The words 'bead' and 'peg' were credited with one point each if used correctly to identify the corresponding item, at least once, during the discourse task. Acceptable synonyms were also given credit, and included words such as, *ball* and *block* for the target vocabulary word *bead,* and the words '*stick* and *stand '* for the target word *peg*. Spatial/Location terms (e.g., here/there, middle, tall) and sequential words (e.g., first, next, then) were credited, with one point each, if the participant used one or more of these during the problem-solving portion of the task.

License