Children's understanding of others' lexical knowledge

Anonymous CogSci submission

Abstract

To communicate successfully, we need to use words that our conversational partner understands. Adults maintain precise models of the words people are likely to know, using both their experience with their conversational partner and general metalinguistic information. Do children also know what words others are likely to know? We asked children ages 4-8 ($n = 62$) to predict whether a very young child would know each of 15 familiar animal words, all typically acquired within a 6-month range. With minimal information, even children as young as 4 made reliable predictions about the target child's lexical knowledge. Further, children's accuracies improved significantly over the course of development. Thus, even preschool age children are adept at inferring other children's vocabulary knowledge, and they could leverage this information to communicate effectively.

Keywords: communication, metalinguistic, knowledge reasoning, cognitive development

Introduction

Imagine visiting the zoo with your friend and their 2-year-old. As you walk by the peacocks, you hear your friend say, "Do you see those blue birds?" Immediately, you know that your friend is talking to their child and not you. If they were talking to you, saying "peacock" would be perfectly clear; however, "blue bird" might be a better description for a child who has never seen a peacock before. Even when talking about the same object, we use different words depending on what we think our conversational partners know and don't know.

Adults track and adapt to their conversational partners' knowledge with relative ease. For example, adults reduce the information they give when re-telling a story to someone who has heard it before, but not when telling the story to a new partner (Galati & Brennan, 2010). Adults can adapt even to partners who are quite different from them, as in the case of parents and their children. Parents model the fine-grained details of their children's vocabularies and use these models in spontaneous communication (e.g. using "blue bird" to describe a peacock, Leung, Tunkel, & Yurovsky, n.d.). These vocabulary models are shaped by extensive individual parent-child interactions, but likely also general metalinguistic knowledge–for instance, that shorter words are typically simpler than longer words and thus learned earlier likely to be known. In line with this conjecture, Kuperman, Stadthagen-Gonzalez, & Brysbaert (2012) asked people to report the age at which they understood a set of 30,000 English words. They found that adults typically overestimated the ages at which words are learned, but were quite accurate at judging their relative order. Thus, even adults without children have access to the kind of metalinguistic information they could use to calibrate speech to children's knowledge.

Can children use this same kind of information to predict whether others will know words? Pre-school age children are able to infer other people's knowledge from relatively sparse information, such as age or expertise (Jaswal & Neely, 2006; Lutz & Keil, 2002). Further, children ascribe different levels of general knowledge to infants, preschool children, and adults (Fitneva, 2010, 2010; Taylor, Cartwright, & Bowden, 1991; VanderBorght & Jaswal, 2009). However, reasoning about another person's specific lexical knowledge may prove difficult for young children. For instance, children often overattribute knowledge to others, especially knowledge they themselves already have (e.g., Gopnik & Astington, 1988; Birch & Bloom, 2003; Ghrear, Fung, Haddock, & Birch, 2020; Taylor, Esbensen, & Bennett, 1994). Children as young as 5 have been shown to make accurate metalinguistic judgments about their own vocabulary knowledge (Walley & Metsala, 1992). Can they use this knowledge to reason about other children as well?

However, reasoning about another person's specific lexical knowledge may prove difficult for young children as they also show consistent errors in reasoning about other's knowledge, commonly over-attributing knowledge (e.g., Gopnik & Astington, 1988; Taylor et al., 1994). The bias to over-attribute knowledge is particularly pronounced when the child themselves knows the piece of information (Birch & Bloom, 2003; Ghrear et al., 2020).

We ask how children infer another children's lexical knowledge, and specifically whether they make word-level predictions in line with expected AoA. Children at least as young as 5 can make metalinguistic judgments about their own vocabulary knowledge, accurately estimating the age and order in which they learned a variety of words (Walley & Metsala, 1992). In our study, we introduced 4- to 8-year-old children to a younger fictional child, and asked them to make judgments about the target child's knowledge of various familiar words. We find that even 4-year-olds make accurate inferences about the target child's knowledge, and that older children's responses more closely correlate with adult judgments of AoA.

We examined the development of children's use of met-

alinguistic lexical knowledge, asking how children's abilities to make word-level knowledge judgments about a younger child change over the preschool and early school years. We found the accuracy of children's judgments improved reliably from 4- to 8-years of age, but that even the youngest children's judgments showed significant knowledge about the words that a younger child was likely to know. Further, 7- and 8-year-old children were more accurate than a comparison group of adults. We also asked children to give a brief explanation for their reasoning, and these responses show that children were most likely to rely on metalinguistic knowledge when making judgments. Our findings show that even young children are adept at estimating other people's specific knowledge, even with minimal information about the other person.

Method

Stimuli   Our stimuli consisted of 15 words drawn from a single domain (animal words), along with corresponding images of each animal. We pulled all animal images ($n = 45$) from a normed image set (Rossion & Pourtois, 2004, recoloring of @snodgrass1980). To ensure our stimuli set spanned a range of ages of acquisition (AoAs), we ranked the animal words from earliest to latest AoA, using data from Kuperman et al. (2012), and split the words into five bins. In order to select animal images that are recognizable and typically identified by a single name, we chose the three animals from each AoA bin with the highest naming agreement according to a naming task with children (Cycowicz, Friedman, Rothstein, & Snodgrass, 1997).

The resulting animals, order by AoA, were dog, duck, cat, pig, fish, turtle, zebra, elephant, snake, penguin, gorilla, owl, raccoon, leopard, and lobster. Although adult AoA estimates for these words range from 2.5 to 7.5 years old (Kuperman et al., 2012), all of these animal words are generally acquired by age 3 according to parent reports of children's vocabulary knowledge (Frank, Braginsky, Yurovsky, & Marchman, 2017). Because the youngest children in our study are 4 years old, we expected all participants to know these animal words.

add thing about why were using kuperman, expand on why it's an okay thing to use, here or results

Participants   We pre-registered a planned sample of 60 children ages 4-8, with 12 children recruited for each year-wise age group. Due to overrecruitment, our final sample included 62 children (12 4-year-olds, 13 5-year-olds, 13 6-year-olds, 12 7-year-olds, 12 8-year-olds). All analyses hold when looking only at the 60 children run first chronologically. Based on a pre-registered exclusion criterion, children who failed to answer all of the questions were excluded and replaced (an additional 7 children). Families were recruited online, primarily through a US University database of families who have expressed interest in doing research or previously participated. Children completed this study over Zoom, interacting with a live experimenter who navigated a slide-style, animated Qualtrics survey.

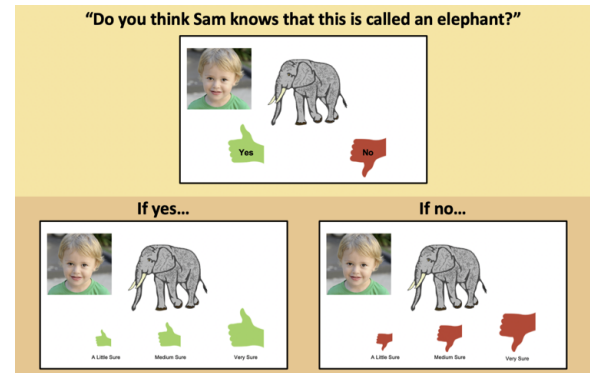A separate sample of 30 adults were recruited via Amazon

Figure 1: The structure of an example trial. The experimenter labeled the animal, then asks the child "Do you think Sam knows that this is called an elephant?" Based on their response, children were then asked to provide a confidence judgment on a 3-point scale (little sure, medium sure, very sure).

Mechanical Turk. The adult sample provides a simple test that our task elicits robust inferences about the target child's lexical knowledge, and that these inferences correspond to extant AoA data. Adult participants completed the same task using Qualtrics, with minor modifications as described below.

Procedure

Introduction. Children were shown a picture of a child named "Sam" (seen in Figure 1). Children were anchored to Sam's knowledge of various familiar skills, specifically some skills that Sam has acquired (e.g., coloring), and some that Sam has not yet acquired (e.g., reading). Children were then specifically anchored to Sam's possible word knowledge in an unrelated domain–given an example of one word Sam knows (*car*), and one word that Sam does npt know (*piano*). This introduction was intended to familiarize the children with Sam, roughly anchor them to Sam's knowledge and age, and to ensure that children understand there are things Sam does not know yet (even things children themselves likely know, such as how to read).

Trial structure. On each trial, children were shown a drawing of a familiar object or animal (Rossion & Pourtois, 2004). The experimenter first labeled the object for the child (e.g., saying "Look, it's [an elephant]!), and then asked the target child's knowledge (e.g., saying,"Do you think Sam knows that this is called [an elephant]? Yes or no?"). Based on their response, children were then asked a follow-up question: "How sure are you that Sam [knows/doesn't know] that this is called [an elephant]–a little sure, medium sure, or very sure?"

All questions were presented with accompanying pictures of thumbs [up/down] of varying size (see Figure 1). Children as young as 3 are able to engage in uncertainty monitoring and report confidence, although these skills do develop in the preschool years (Lyons & Ghetti, 2011). Children's responses to these two items were recoded onto a 1-6 scale from 1–*very sure Sam doesn't know* to 6–*very sure Sam knows*.

The experimenter provided no evaluative feedback on any trials, but did offer consistent neutral feedback (e.g., repeating the child's answer or saying "Okay!"). When a child failed to respond within about 5 seconds or offered a non-canonical response (e.g., saying "Maybe"), the experimenter acknowledged the child's answer and then repeated the question with the possible responses. If a child did not answer after the question was repeated, the experimenter moved on and marked the trial as no response. These were considered "incomplete" sessions and these participants were not included in the final sample.

Familiarization trials. Children first completed two non-animal familiarization trials, one for an early-acquired word (ball) and one of a late-acquired word (artichoke). These trials followed the trial structure described above and were intended to help familiarize children with the structure of the questions and scales. These trials were always asked first and in a fixed order.

Animal trials. Children were then shown 15 trials of the same form (see example trial in Figure 1). For the 15 animal trials, trial order was randomized across participants to control for any potential order effects in children's responses.

Explanation. After completing the final animal trial, children were asked an open-ended explanation question about their final judgment (e.g., "Why do you think Sam [knows/doesn't know] that this is called [an elephant]?"). Because the trial order was randomized, the explanations concerned different animal words across participants.

Final check questions. Children were asked two questions about Sam's skill knowledge, one early acquired skill (going up and down stairs) and one very late acquired skill (driving a car). These questions again followed the general trial structure described above. The skill knowledge items were included as an additional check that children at all ages were able to use the scale appropriately, in case young children failed to differentiate animal words based on AoA. Lastly, children were asked to report how old they thought Sam was. This question was intended to assess another aspect of children's belief about Sam. Sam's photo and skill knowledge were intended to indicate toddlerhood.

Adult procedure. Adult participants completed a minimally adapted version of the same task online via Qualtrics. Unlike children, adults were simply presented with the full 6-point scale (1–*very sure Sam doesn't know* to 6–*very sure Sam does know*). Additionally, the task was administered asynchronously, so adult participants did not interact with an experimenter or receive any feedback during the task. Other-
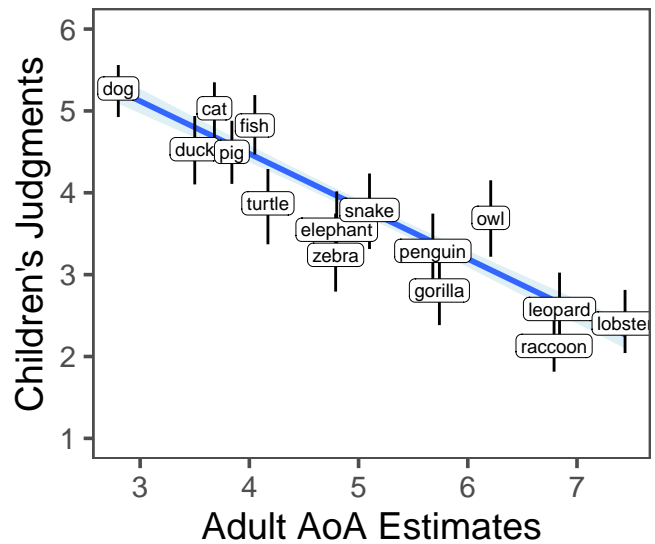

Figure 2: Comparing adult AoA estimates (in years, taken from Kuperman et al., 2012) and children's judgments on our 6-point scale (1 = very sure Sam doesn't know; 6 = very sure Sam knows). The black lines show 95% confidence intervals for each item. The shaded region shows one standard error based on a linear regression estimated from the raw data.

wise, the adult task was identical to the child task described above.

Results

Our primary analyses compare knowledge judgments on our 6-point scale to AoA judgments from adults (taken from Kuperman et al., 2012). Data were analyzed using pre-registered mixed effects model predicting children's judgments from adult AoA estimates (Kuperman et al., 2012), including random effects for participant and word. Using the lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015). We our model syntax was judgment ~ AoA + (1 | participant) + (1 | word).

We expected that overall, children's judgments would recover the ordinal shape of age of acquisition data for these items. That is, children would infer that the target child is most likely to know early acquired words, and least likely to know late acquired words. As a result, we expected a negative relationship between judgments of the target child's lexical knowledge and adult AoA estimates.

First, analyzing adults responses on our task, we saw the predicted negative effect of AoA on adults' judgments of the target child's knowledge (see Figure 3, $\beta = -0.63$, $t = -8.71$, $p < .01$). This confirms that our task is eliciting reliable predictions from adults, and that adults' inferences about the target child's knowledge match predictions from AoA estimation tasks (Kuperman et al., 2012).

Do children's judgments about a child's vocabulary knowledge also reflect a sensitivity to which words are learned
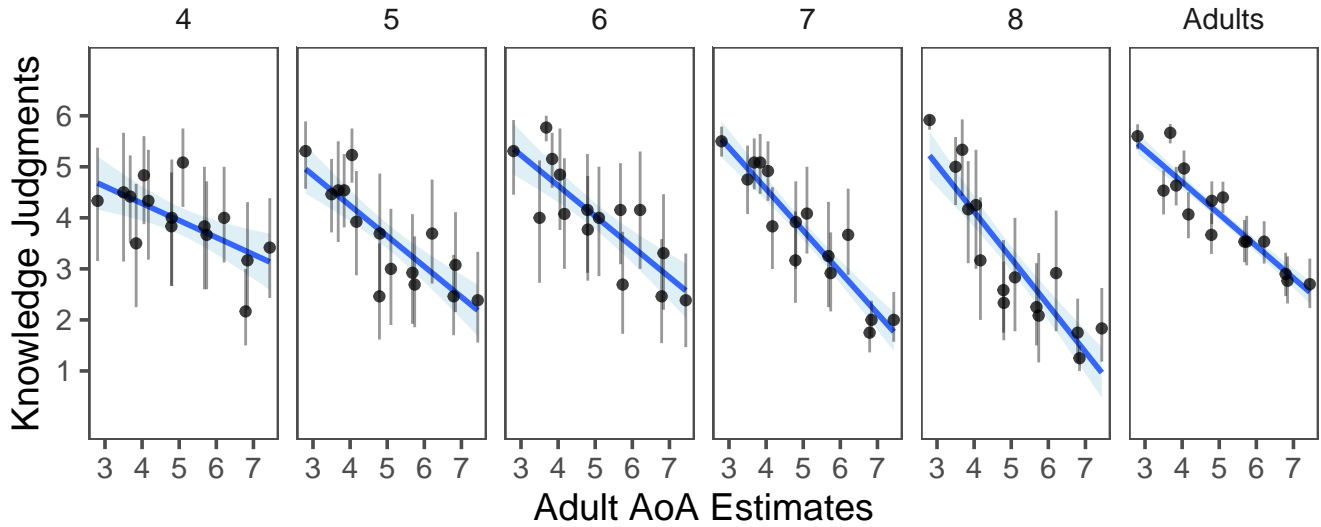
Figure 3: Children and adults judgements about the target child's word knowledge across development, compared with adult AoA estimates (in years, taken from Kuperman et al., 2012). Each point represents 1 of the 15 word items, with black lines showing 95% percent confidence intervals for each item. The shaded region shows one standard deviation based on a linear regression estimated from the raw data.

later? To answer this question, we ran a mixed effects model: `judgment ~ AoA + (1 | participant) + (1|word)`. We find a significant negative effect of AoA on children's judgments ($\beta = -0.65$, $t = -8.29$, $p < .001$). That is, overall, children judged that the target child would be most likely to know an early acquired word (e.g., dog) and least likely to know a late acquired word (e.g., lobster, see Figure 2).

We also expected developmental change in children's sensitivity to Sam's vocabulary knowledge, with older children's judgments recovering word-level AoA data more closely. To test for developmental changes in children's responses, we used the same mixed effects model but included an effect of age and an interaction between AoA and age. Our model syntax was `judgment ~ AoA * age + (1 | participant) + (1|word)`. We expected a significant interaction between AoA and child;s age, consistent with older children/s judgments more closely reflecting word-level AoA data. That is, when plotting children's judgments against adult AoA estimates, older children would show steeper negative slopes than younger children (Figure 3). As above, our model shows the same main effect of AoA that we saw in the overall model ($\beta = -0.65$, $t = -8.29$, $p < .001$). We also see a positive main effect of children's age on their ratings ($\beta = 0.55$, $t = 3.68$, $p < .001$). Crucially, we see the expected interaction between child's age and adult's estimated AoA ($\beta = -0.14$, $t = -5.1$, $p < .001$), suggesting that children's judgments are becoming more adult-like in this age range (Figure 3).
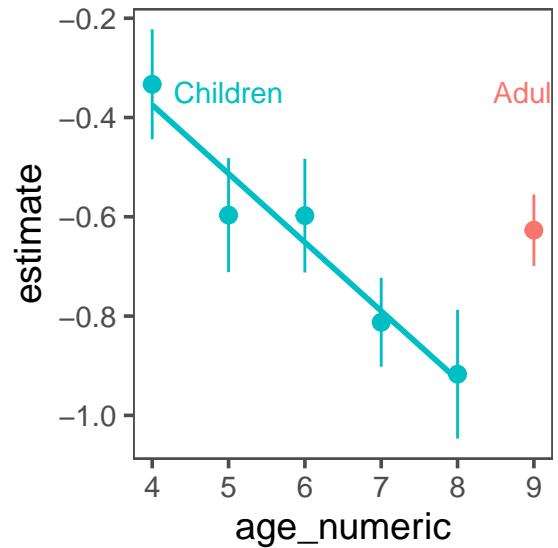
To test the robustness of this intuition at each age, we ran the above model separately for each year-wise age group. While we see evidence of developmental change above, this additional analysis helps us understand if even young children are showing this intuition. We find a significant negative effect of AoA on children's judgments at all age groups (with the smallest effect in 4-year-olds: $\beta = -0.33$, $t = -3.01$, $p = .01$). That is, even 4-year-old children judged that late-acquired animal words were less likely to be known by the target child.

Secondary Results

Familiarization trials. Children completed two familiarization trials asking about the target child's knowledge of two non-animal words (*ball* and *artichoke*). These questions were primarily included to help children get accustomed to the general trial structure. To analyze children's responses on these familiarization items, we used a similar mixed effects structure predicting children's knowledge judgments from the item and a random effect of pariticpant.

Overall, children were significantly more likely to report that Sam knows the word "ball" (*mean* = 4.65) than that Sam knows the word "artichoke" (*mean* = 1.87, $\beta = 2.77$, $t = 10.6$, $p = < .01$). Analyzing judgments separately for each age group, 4-year-olds do not significantly differentiate the two familiarization items ($\beta = 0.17$, $t = 0.21$, $p = .83$). All other age groups signifcantly differentiated the two familiarization items ($ps < 0.05$).

Skill knowledge. We included two questions about the target child's skill knowledge. In the event that children at any age were differentiating the target child's lexical knowledge, these questions were intended primarily as a check that children at all ages were nevertheless able to use the scale appropriately and able to infer knowledge in another domain. Note that the two skill items (*going up and down stairs*, and *driving a car*) are also in line with children's own knowledge, unlike the word items. That is, child participants should be able to answer these questions appropriately even if they are reasoning egocentrically about their own knowledge.

Overall, children differentiated the target child's skill knowledge on these two items. We used a similar mixed effects structure predicting children's knowledge judgments from the item and a random effect of pariticpant. Children were significantly more likely to report that the child knows how to go up and down stairs (*mean* = 4.1) than that the child knows how to drive a car (*mean* = 1.4, $\beta = 2.69$, $t = 10.29$, $p = < .01$). Analyzing judgments separately for each age group, even 4-year-olds were significantly differentiating the two skill item ($\beta = 2.33$, $t = 4.31$, $p = < .01$).

Explanations As a secondary analysis, we were also interested in the reasons young children gave for why the target child would or would not know a given word. While children sometimes offered spontaneous explanations throughout the study, this analysis focuses on the explanation elicited after the final animal trial. Based on preliminary discussions between the authors, the explanations were divided into 6 non-mutually-exclusive categories: *Language*, *Experience*, *Location*, *Age*, *Unsure*, and *Other*.

Language reflects explanations that explicitly appealed to language properties. *Experience* reflects explanations that appeal to the child's real-world experience with the referent. *Location* reflects explanations that specifically reference a particular place the animal is associated with. *Age* reflects explanations that reference a particular age or general age-
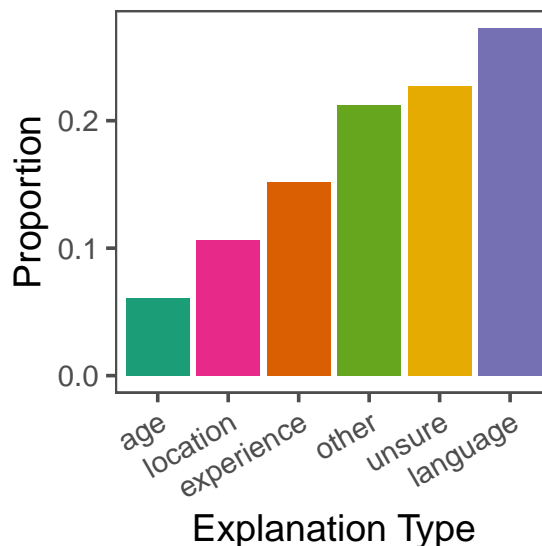

Figure 4: Children's explanations for why they think the target child knew or didn't know an animal. Categories are not mutually exclusive.

group. Any child that failed to answer the explanation question or expressed ignorance was coded as giving an explanation of *Unsure*. An explanation that didn't fall into any of the above category was coded as *Other*. Note that coding was not mutually-exclusive, so explanations could be coded as including multiple categories. See Table 1 for examples of each coding category. Figure 4 shows the proportion of children who gave each type of explanation.

To understand how children's explanations may change over development, we looked at explanations by age group. *Language* explanations were used by the highest proportion of children overall (27.27%). Do older children account for all of those explanations? While almost half of the 7- and 8-year-olds appealed to *Language* explanations (41.67% and 45.45%), we find that 5-year-olds (but not 6-year-olds) were also likely to use *Language* explanations (38.46%). Unsurprisingly, *Unsure* explanations were more common in younger children, accounting for over half (58.33%) of the 4-year-olds' explanations, and none of the 7- and 8-year-olds' responses.

Discussion

Our ability to infer other people's knowledge is crucial for successful communication. Young children are capable of inferring others' general knowledge states, but do they make accurate judgments about another person's specific knowledge? We asked 4- to 8-year-old children to estimate another child's knowledge of animal words, and found that children as young as 4 are sensitive to a younger child's vocabulary knowledge. Children across age groups made judgments similar to those of adults, with older children recovering more adult-like patterns.

Category	Example Utterance
Language	Because it was a very long word.
	Because it only has 3 letters.
Experience	Because maybe he has a dog.
	Because gorillas are really rare animals
Location	Because penguins live in the artic and it's too cold for little kids...
	Because fish swim under the ocean.
Age	Because I think I knew that when I was around 3...
	Because if he went to preschool then he probably knew it...
Unsure	I don't know.
	I'm not sure.
Other	Because it had a longer beak than a bird.
	Because it's small.

Table 1: Example explanations from child participants for each of the five categories used for coding.

Our findings indicate that young children have surprising metalinguistic knowledge, and can use that knowledge to make highly-specific inferences about other people's knowledge. The animal words used in our study are generally learned within a 6-month period, yet young children still distinguished early-acquired words from late-acquired words in this set. Our study also builds upon the extant literature on children's inferences about other people's knowledge to show that children infer others' specific, lexical knowledge. When given fairly minimal information about another child, children readily make estimates about that child's lexical knowledge.

Surprisingly, responses from the oldest children in our sample more closely correlated with adult judgments of AoA (Kuperman et al., 2012) even when compared to adults participants in this study. Older children in our study were more likely to use the full 6-point scale. A potential explanation is that children may be more accurate at estimating other chilren's lexical knowledge. In line with this hypothesis, Walley & Metsala (1992) found that preschoolers were better at estimating the age at which they learned certain words than adults were. One reason for this may be that adults cannot rely on specific memories for when they learned words, and must use other information such as word frequency. However, this is unlikely to be the complete story, as this account would predict that younger children would perform better than older children when making inferences about another young child, which is the opposite of our findings. Additionally, children and adults' estimates for how old the target child was may be an important factor. In our sample, adults gave higher estimates for the target child's age when compared to children. This may, in part, explain why adults' inferences did not entirely match Kuperman et al. (2012) AoA data, and why they were less likely to say that the target child did not know a certain animal word.

How are children in our study making estimates about other people's knowledge? Children's own explanations suggest that they use various cues to make their estimates. Overall, language-related explanations were used by the highest proportion of children, and even preschool age children appealed to this explanation. This result suggests that even young children rely on metalinguistic knowledge to make inferences about others. One limitation of the current study is that it leaves the mechanisms underlying such estimates unclear. Future work should more directly probe the features underlying this inference– to see if children are relying on their own uncertainty, word length (and other linguistic cues), features of the referent itself, or still other features.

The current work lays the foundation for future research on how children leverage their knowledge of other people to communicate successfully. While some studies have found that young children struggle in a variety of communicative tasks (e.g. Krauss & Glucksberg, 1977), other work has shown that by age 5, children selectively talk about general or specific characteristics of an object based on their partners' knowledge state (Baer & Friedman, 2018). Why might children struggle in some situations and not others? Our work can begin to address this by mapping out whether communicative difficulties stem from tracking an interlocutor's knowledge, or problems using that information in language production.

References

Baer, C., & Friedman, O. (2018). Fitting the message to the listener: Children selectively mention general and specific facts. *Child Development*, *89*(2), 461–475.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. http://doi.org/10.18637/jss.v067.i01

Birch, S. A., & Bloom, P. (2003). Children are cursed: An asymmetric bias in mental-state attribution. *Psychological Science*, *14*(3), 283–286.

Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: Norms for

name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology*, *65*(2), 171–237.

Fitneva, S. A. (2010). Children's representation of child and adult knowledge. *Journal of Cognition and Development*, *11*(4), 458–484.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, *44*(3), 677–694.

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, *62*(1), 35–51.

Ghrear, S., Fung, K., Haddock, T., & Birch, S. A. (2020). Only familiar information is a "curse": Children's ability to predict what their peers know. *Child Development*.

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, 26–37.

Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science*.

Krauss, R. M., & Glucksberg, S. (1977). Social and nonsocial speech. *Scientific American*, *236*(2), 100–105.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 english words. *Behavior Research Methods*, *44*(4), 978–990.

Leung, A., Tunkel, A., & Yurovsky, D. (n.d.). Parents fine-tune their speech to children's vocabulary knowledge. *Psychological Science*.

Lutz, D. J., & Keil, F. C. (2002). Early understanding of the division of cognitive labor. *Child Development*, *73*(4), 1073–1084.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33*, 217–236.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174.

Taylor, M., Cartwright, B. S., & Bowden, T. (1991). Perspective taking and theory of mind: Do children predict interpretive diversity as a function of differences in observers' knowledge? *Child Development*, *62*(6), 1334–1351.

Taylor, M., Esbensen, B. M., & Bennett, R. T. (1994). Children's understanding of knowledge acquisition: The tendency for children to report that they have always known what they have just learned. *Child Development*, *65*(6), 1581–1604.

VanderBorght, M., & Jaswal, V. K. (2009). Who knows best? Preschoolers sometimes prefer child informants over adult informants. *Infant and Child Development: An International Journal of Research and Practice*, *18*(1), 61–71.

Walley, A. C., & Metsala, J. L. (1992). Young children's age-of-acquisition estimates for spoken words. *Memory & Cognition*, *20*(2), 171–182.