Metaphor comprehension in preschool children

Ariel Starr
Taylor Petersen
Sahana Sridhar
University of Washington, USA

Abstract: Recent work suggests that though young children can comprehend metaphors based on shared perceptual or functional features of objects, comprehending metaphors based on abstract relations across domains presents a greater challenge. We conducted two pre-registered studies (n = 272; mean age = 3.77 years; 143 female) to investigate children's ability to understand metaphors based on object and abstract similarities. We also assessed how children's language learning environments (monolingual or bilingual) relate to their metaphor comprehension. Children were successful in understanding both types of metaphors. In addition, monolingual and bilingual children were equally proficient in metaphor comprehension. These findings highlight the sophisticated ways that preschool-aged children can use their rapidly developing lexicons.

Keywords: metaphor comprehension, bilingualism, language development

Corresponding author: Ariel Starr, Department of Psychology, University of Washington, Seattle, WA 98195, USA. Email: abstarr@uw.edu.

ORCID ID(s): https://orcid.org/0000-0003-0433-5003 (Starr, A.); https://orcid.org/0009-0000-6272-0191 (Sridhar, S.)

Citation: Starr, A., Petersen, T., & Sridhar, S. (2025). Metaphor comprehension in preschool children. *Language Development Research*, 5(2), 1–27. http://doi.org/10.34842/starr

Introduction

Metaphors are a powerful tool for expression. By highlighting similarities between unrelated domains, metaphors drive creativity in artistic and scientific fields and shape how we think about the world around us (Bowdle & Gentner, 2005; Kuhn, 1993; Lakoff & Johnson, 1980). However, understanding metaphors requires understanding figurative language, which can be difficult for young children. Adults can rest reasonably assured that when someone is described as a "busy bee," there are no actual bees present. To young children, however, this statement might cause confusion (Vosniadou, 1987). How do children come to be able to understand, and even appreciate, metaphors?

A classic view has been that preschool-aged children do not understand metaphors and instead interpret all language literally (e.g., Silberstein et al., 1982; Winner et al., 1976, 1980). However, as many have noted (e.g., Gentner, 1988; Pouscoulous, 2011; Vosniadou et al., 1984; Winner et al., 1976), the literature is rife with examples of children seemingly producing metaphorical language (e.g., an 18-month-old referring to a toy car as a "snake" while twisting it up and down his mother's arm (Winner et al., 1979)). While a more charitable interpretation of these observations might be that young children produce metaphors before they are able to comprehend metaphorical languages produced by others, these instances were often interpreted as overextension or pretend play rather than use of metaphorical language. However, early studies used metaphorical language and concepts that are unlikely to be familiar to young children (Winner et al., 1976; see Winner et al., 1980 for a detailed discussion of metaphor task demands). As a result, these studies may have mischaracterized children's understanding of metaphorical language (Vosniadou, 1987). When young children are presented with metaphors that use familiar language and concepts, laboratory studies do reveal early metaphoric competence. For example, by age 5, children can use toys to enact metaphorical phrases like "Billy was a squirrel burying the nuts" (correctly interpreted as meaning that Billy was hiding cookies rather that literally burying nuts; Vosniadou et al., 1984) and provide explanations of metaphorical phrases like "plant stems are drinking straws" (Gentner, 1988).

In addition to ensuring that concepts are age-appropriate, investigating children's metaphorical understanding also requires recognizing that metaphors can take many different forms. Gentner (1988) argued that metaphors can be divided into the partially overlapping categories of attributional metaphors (those based on perceptual similarities), relational metaphors (those based on a common relational structure, such as function), double metaphors (those that involve both perceptual and relational similarities), and complex metaphors (those that are idiosyncratic). Gentner found that children as young as 5 were able to understand attributional metaphors, and a preference for relational metaphors increased with age (1988). These results let to the hypothesis that the ability to understand metaphors based on shared perceptual properties emerges earlier than the ability to understand metaphors based on shared relations.

Recent studies using simplified paradigms have demonstrated that even preschool-aged children can understand metaphorical language when it involves familiar object concepts (Pouscoulous & Tomasello, 2020; Zhu et al., 2024; Zhu & Gopnik, 2023).In Pouscoulous and Tomasello (2020), for example, 3-year-old participants were presented with two toys, and an experimenter asked the child for one of the toys using a novel attributional metaphorical phrase structured in the form of "the X with the Y," where X refers to the object and Y refers to a figurative perceptual attribute. The metaphorical phrase "the dog with the brown shoes," for instance, referred to a toy dog with brown feet. In this simplified task, 3-year-olds succeeded in choosing the toys that corresponded to the metaphorical phrases. The metaphorical phrases used by Pouscoulous and Tomasello (2020) involved metaphors that evoked shared perceptual features and common knowledge about objects. In the example metaphor "the dog with the brown shoes," children can use their knowledge of shoes (i.e., they go on feet) to solve the metaphor and match the phrase to a picture of a dog with brown feet. These results therefore extend Genter's 1998 findings down to younger children and demonstrate that the ability to understand attributional metaphors is present by age three.

In addition to succeeding on attributional metaphors based on perceptual similarities, recent studies have shown that preschoolers can also comprehend relational metaphors that are based on shared relations between objects, namely object functions (Zhu et al., 2024; Zhu & Gopnik, 2023). Zhu and colleagues (2024) found that children as young as 3 could understand relational metaphors that make comparisons between attributes of two objects that serve the same purpose (e.g., "pools are bowls," which makes use of the fact that both pools and bowls hold liquids). Three-year-olds could also use their object knowledge to make novel metaphorical inferences (Zhu & Gopnik, 2023). For example, after hearing the phrase "Blickets are eyes," 3-year-olds decided that blickets would be more likely to help you see things than help you chop things. These recent studies demonstrate that preschoolers are adept at comprehending metaphors based on attributional and functional features of familiar objects.
In contrast, young children have more difficulty with metaphors that involve abstract properties without direct physical or perceptual correlates, such as the metaphorical motion of time (e.g., "time *flew* by") (Özçalışkan, 2005; Özçalışkan, 2007; Stites & Özçalışkan, 2012). Özçalışkan found that while 3-year-old children performed at chance in understanding these abstract metaphors, 4- and 5-year-old children understood their meanings. By age 5, children could also provide verbal explanations of the concepts represented in metaphorical motion. Taken together, the full picture of prior work on metaphor comprehension in preschool-aged children suggests that comprehension may development gradually, with understanding of metaphors based on perceptual and functional similarities developing earlier than understanding of

abstract metaphors (Gentner, 1988; Vosniadou, 1987).

The main goal of the present study was to further investigate preschoolers' ability to comprehend abstract metaphors. We use the term abstract metaphor to refer to metaphors that involve a conceptual mapping between two unrelated domains (Casasanto, 2009; Lakoff & Johnson, 1980) instead of highlighting a shared perceptual or functional feature of a physical object. An example of an abstract metaphor is "the girl is feeling *down*," which involves mapping emotional states to spatial positions. When making sense of metaphors based on perceptual or functional similarities, children are making mappings between physical objects (e.g., shoes and feet). By contrast, abstract metaphors require making associations between two unrelated concepts (e.g., spatial locations and emotional valence) or mapping cross-modally across different sensory perceptions (e.g., mapping visual brightness to auditory volume, as in describing a bright color as *loud*). Although prior work by Özçalışkan and colleagues (2005, 2007, 2012) has indicated that abstract metaphors can be difficult for young children, we attempted to facilitate preschoolers' understanding by providing visual illustrations to go with the verbal metaphors. With this simplified presentation, we predicted that preschool-aged children would be able to understand abstract metaphors that involved familiar words and concepts, though they might find these metaphors more difficult than metaphors based on object knowledge.

A second goal of the present study was to investigate how children's language learning environments influence the development of metaphor comprehension skills. In addition to pragmatic and conceptual constraints that contribute to children's difficulties with metaphorical language, metaphorical language may also be difficult for young children to understand because they are resistant to assigning multiple labels to a single concept (Rubio-Fernández & Grassmann, 2016). In their study, Rubio-Fernández and Grassmann found that 3- and 4-year-old children were more successful at choosing the object pair that matched a metaphorical phrase when the target of the metaphor was a toy block compared to when it was a familiar object (e.g., "the train with a hat" with the hat represented by either a block or a toy dog). Children's difficulty with metaphorically extending a word to an object with a known label (e.g., referring to a dog as a hat) may stem from the mutual exclusivity heuristic. When young children are learning new words, they frequently operate under the heuristic that each concept has a single label and therefore, if a concept already has a known label, a newly introduced word likely labels a new concept (Markman & Wachtel, 1998). The mutual exclusivity heuristic has been documented in children as young as 17 months, who will preferentially look at a novel object compared to a familiar object when they are presented with a novel object label (Halberda, 2003).

However, children's reliance on mutual exclusivity and their expectation that there is a one-to-one mapping between labels and concepts is shaped by their language learning environment (Houston-Price et al., 2010). Children who are learning multiple languages are frequently learning multiple labels for the same concepts. As might be expected based on their learning history, bilingual toddlers are less likely than monolingual toddlers to assume that a novel label refers to a novel object rather than a familiar object (Houston-Price et al., 2010; Repnik et al., 2021). Similarly, bilingual preschoolers are less likely to use mutual exclusivity to learn new adjectives in comparison to monolingual preschoolers (Groba et al., 2019), which demonstrates that the phenomenon is not specific to learning nouns. Some studies have also suggested that bilingual children have stronger metalinguistic skills in comparison to monolingual skills (Ben-Zeev, 1977; Diaz & Farrar, 2018). Metalinguistic awareness includes a range of abilities including phonological awareness and understanding of homonyms and synonyms. Understanding homonyms and synonyms requires children to flexibly switch between different meanings of a word depending on the context – an ability that is also relevant for understanding metaphors. Because bilingual children are more open to the possibility that known concepts can have more than one label and labels can refer to multiple concepts, they might have an advantage in understanding metaphorical language. Although rates of bilingualism are increasing in the US and around the world (Luk, 2017), studies on metaphor comprehension in children have focused on monolingual children. To better understand different factors that influence the development of metaphor comprehension, the present study included monolingual and bilingual children.

In the present study, we compared monolingual and bilingual preschoolers' comprehension of two types of metaphors. Prior studies have shown that preschool-aged children can understand metaphors that involve shared perceptual and functional features of objects. We compared children's performance on these types of metaphors, which we term object metaphors, with performance on metaphors we term abstract metaphors, metaphors that involve children flexibly extending words across modalities and perceptual domains. We predicted that children would be able to understand both types of metaphors, but that they would be more successful in understanding perceptual metaphors compared to abstract metaphors. The second goal of the study was to explore how children's early language learning environments affect metaphor comprehension. We predicted that because of their regular exposure to second labels, bilingual children might be more successful in understanding both types of metaphorical language in comparison to monolingual children.

We tested these hypotheses in two pre-registered studies that used a metaphor comprehension paradigm adapted from Pouscoulous and Tomasello (2020). This paradigm involves matching a metaphorical phrase to one of two pictures, which may be an easier task for preschool-aged children than comprehending abstract metaphors presented in speech alone, as in Özçalışkan (2005; 2007). Children were also asked to select pictures matching the literal meaning of the target words. The vocabularies of young children typically contain more words for concrete nouns relative to words for more abstract concepts and more nouns in general relative to adjectives and verbs (L.

Bloom et al., 1975; Braginsky et al., 2019; Casasanto et al., 2010; Gentner, 1978). Furthermore, children's understanding of concrete word meanings precedes their understanding of more abstract meanings (Andrews & Madeira, 1977; P. Bloom, 2000; Braginsky et al., 2019; Casasanto et al., 2010; Clark, 1973; Gentner & Boroditsky, 2001). We predicted that children would need to understand the literal meaning of words in order to extend them metaphorically. Prior work has implicitly built off this assumption by testing children's knowledge of the literal meanings of the target word stimuli and ensuring that children understood the majority of the words (Pouscoulous & Tomasello, 2020), assuming that children would know the meaning of the target words without explicitly testing (Zhu et al., 2024), or circumnavigating this issue by using novel words as targets (Zhu & Gopnik, 2023). In Experiment 1, we opted to test children's understanding of the literal meaning of words and set a threshold performance level to include children in our analyses, as in Pouscoulous and Tomasello (2020). However, this method involved analyzing some metaphor trials for which children had not succeeded in identifying the literal meaning of the target word. In Experiment 2, we opted to only analyze metaphor comprehension trials for which children answered the corresponding vocabulary question correctly. Experiment 1 was conducted remotely with an experimenter present. Experiment 2 was a conceptual replication to test the robustness of the Experiment 1 results and was conducted remotely in an unmoderated format.

Study 1

Method

Participants

Data from 123 participants aged 2.5- to 4-years-old (mean = 3.6 years, range = 2.5-4.8, 64 female) were included in the final analyses. Participants were classified as monolingual or bilingual based on their language learning environments as described by their caregivers. Parents were asked if their child was regularly exposed to another language besides English, and if the answer was yes, to estimate the proportion of time the child was exposed to each language. Participants who were exposed to English 90% of the time or more were considered monolingual (n = 74) (Byers-Heinlein, 2017; Jardak & Byers-Heinlein, 2019). Participants who were exposed to English and a second language at least 20% of the time were considered bilingual (n = 49; see Supplement for list of second languages spoken by participants and for the language exposure survey). We preregistered a sample size of 50 monolingual children and 50 bilingual children, but slightly more monolingual children were included in the analyses as we continued recruiting participants until we had reached 50 bilingual children. However, after data collection was stopped, we realized that one child was initially incorrectly classified as bilingual. Data from an additional 70 participants were excluded for performing below 70% on the vocabulary test (n = 34), not fitting into one of our language categories (e.g., regularly exposed to more than two languages;

n = 14), caregiver interference (n = 5), consistently choosing one character over the other (n = 4), parent-reported developmental disabilities (n = 3), distractedness (n = 3), performing below 66% on the attention checks (n = 2), unclear parental language exposure reports (n = 2), failing to complete the experiment (n = 2) or for poor internet connection during the session (n = 1). The University of Washington Institutional Review Board approved this study. Parents of the participants consented for their child's participation in the study and consented for video and audio recording of the study. Parents of participants were compensated with a $5 electronic Tango gift card.

Out of all families who participated in Experiment 1, 160 filled out our optional demographic questionnaire. Because demographics information was collected anonymously, we are unable to report the demographics separately for our monolingual and bilingual groups. Parents identified participants as 47% White, 26% multiracial, 16% Asian, 4% Hispanic or Latine, 0.6% Black or African American, 0.6% American Indian or Alaska Native, 3% another option not listed, and 2.5% chose not to respond. With respect to annual household income, 43% of caregivers reported less than $100,000, 40% reported $100,000-$200,000, 2% reported more than $200,000, and 15% chose not to respond. Data were collected between December 2020 and October 2021.

Procedure

Children were tested by an experimenter over Zoom while their parent sat nearby. Parents were asked to use a laptop or desktop computer to ensure that the stimuli were displayed similarly across participants, and each parent was sent a Zoom user guide prior to the study date. During the experimental session, children played a game in which they chose which of two pictures best matched a metaphorical expression provided by the experimenter. After children completed the experiment, parents completed a questionnaire about their child's language environment. Altogether, the session took approximately 15 minutes.

Materials and Design

Because it is difficult to discern the referent of children's pointing responses over Zoom, each target picture had a small picture of either Cookie Monster or Elmo in a top corner, and children were asked to verbally indicate whether Cookie Monster or Elmo had the picture that best matched the metaphorical phrase. The game began with two warm-up questions to familiarize children with the procedure. The experimenter asked the participant, "Whose picture has a red dot, Elmo or Cookie Monster?" and one character had a red dot below them whereas the other character did not. For the practice questions, the experimenter corrected the child if they were incorrect, but only neutral feedback was provided for the test questions. Children also completed three attention check questions interspersed with the test questions that did not involve metaphorical language (e.g., "Whose picture has the chair?"). Children needed to answer at least two out of three attention check questions correctly for their data to be included in the final analyses.

We created ten novel metaphors that involved concepts and language familiar to preschool-aged children. Five of the metaphors were classified as object metaphors, which were based on shared, visually perceptible features of objects (e.g., a java jacket for "the cup with the sweater") and were similar to those used in Pouscoulous & Tomasello (2020). The other five metaphors were classified as abstract metaphors, which were based on cross-modal mappings rather than physical features (e.g., a girl frowning for "the girl is feeling down"). Participants answered all ten metaphor questions in a randomized order. There were two versions of the experiment, A and B, which were identical except for the abstract metaphor questions involved opposite pairs of words. For example, version A asked about,"a girl who is feeling down," and version B asked about,"a girl who is feeling up." Participants were randomly assigned to a version. See Figure 1 for examples of the stimuli and Table 1 for a list of all metaphors. The mean ages of acquisition (Kuperman et al., 2012) for the target object and abstract metaphor words were not statistically different ($t(10.7) = 0.42$, $p = .68$). Adult English-speaking participants (N = 23) completed a pilot experiment to confirm that the novel metaphors were reliably associated with the target images. This pilot experiment used the same format and questions as the child experiment, except that participants read the questions on their own without an experimenter present. Adult participants chose the target image 99% of the time.

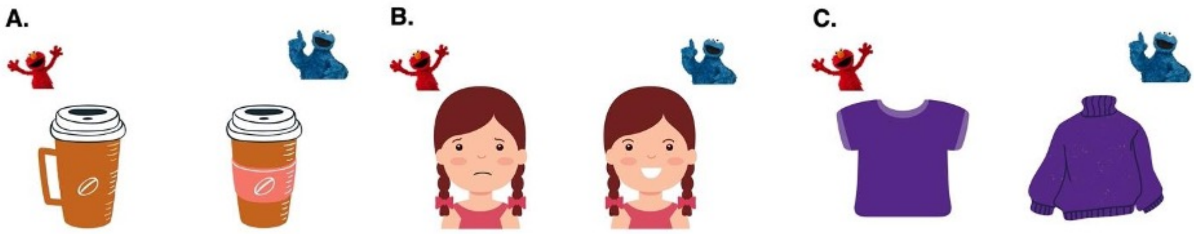

Figure 1. ***Example stimuli for the different types of experimental questions. (A) Object metaphor stimuli for the question "whose picture has the cup with a sweater, Elmo or Cookie Monster?" (B) Abstract metaphor stimuli for the question "whose picture has a girl who is feeling down, Elmo or Cookie Monster?" (C) Vocabulary stimuli for the question "whose picture has the sweater, Elmo or Cookie Monster?***

After the metaphor questions, children completed ten vocabulary questions that assessed their knowledge of the literal meaning of the words involved in the metaphor questions (e.g., "sweater" and "down" from the prior examples). These questions had the same format as the metaphorical questions (e.g., "Whose picture has the sweater?" with response options of a sweater versus a t-shirt). Because we assumed that children needed to understand the literal meaning of a word to be able to metaphorically extend it, and because children may also guess randomly if they are not engaged with the game, only children who selected the correct picture for at least eight out of the ten vocabulary questions were included in the final analyses. For all metaphor and vocabulary questions, only neutral feedback (e.g., "thank you!") was provided.

Table 1. ***Mean metaphor and vocabulary and comprehension accuracy by item.***

Question "Which picture has the..."	**Version**	**Question Type**	**Mean Metaphor Accuracy [95% CI]**	**Mean Vocabulary Accuracy [95% CI]**
Bottle with a big belly		Object	0.75 [0.64-0.86]	0.90 [0.77-1]
Cup with a sweater		Object	0.83 [0.71-0.95]	0.88 [0.75-1]
Dog with socks		Object	0.80 [0.68-0.92]	0.98 [0.83-1]
Horse with a backpack		Object	0.76 [0.65-0.87]	0.98 [0.83-1]
House with eyes*		Object	0.62 [0.53-0.71]	0.97 [0.83-1]
Bird that is happier*	A	Abstract	0.55 [0.44-0.66]	0.98 [0.78-1]
Bird that is sadder*	B	Abstract	0.56 [0.44-0.68]	0.95 [0.75-1]
Boy who is having a bumpy day	A	Abstract	0.77 [0.61-0.93]	0.91 [0.72-1]
Boy who is having a smooth day	B	Abstract	0.66 [0.52-0.80]	0.76 [0.60-0.92]
Dress that is loud*	A	Abstract	0.62 [0.49-0.75]	0.83 [0.66-1]
Dress that is quiet*	B	Abstract	0.54 [0.42-0.66]	0.86 [0.68-1]
Girl who is feeling down*	A	Abstract	0.72 [0.57-0.87]	0.95 [0.75-1]
Girl who is feeling up*	B	Abstract	0.59 [0.46-0.72]	0.92 [0.72-1]
TV that is sick*	A	Abstract	0.58 [0.46-0.70]	0.97 [0.77-1]
TV that is healthy*	B	Abstract	0.58 [0.46-0.70]	0.93 [0.73-1]

Note. Asterisks denote questions with low performance that were replaced in Experiment 2.

Data Analysis

The preregistration and all experimental materials, data, and code required to reproduce the analyses for this study are publicly available at https://osf.io/ta8vf. Our final models deviated slightly from the preregistered models in that models with the full random effects structure did not converge, so we used models with a simplified random effects structure. Data cleaning, analyses, and visualizations were performed in R using the tidyverse and lmerTest packages (Kuznetsova et al., 2017; Wickham, 2014).

Results

Main Analyses

To determine how metaphor type and language learning environment influence children's metaphor comprehension, we ran a pre-registered logistic mixed effects model predicting trial accuracy with metaphor type (object or abstract), language learning environment (monolingual or bilingual), the interaction between metaphor type and language group, and question version (A or B) as fixed effects, along with random intercepts for subject and item. We found a borderline effect of metaphor type (B = 0.49, p = .058) but no significant effect of language environment (B = -0.32, p = .085) or interaction between metaphor type and language environment (B = 0.29, p = .26). Children matched the metaphorical phrase to the correct picture at above chance levels for both types of metaphors (ts > 6.1, ps < 0.001). Performance on object metaphor questions was higher than performance on abstract metaphor questions (object: M = 0.75, SE = 0.02; abstract: M = 0.62, SE = 0.02; Figure 2A). In addition, bilingual children demonstrated descriptively higher metaphor comprehension than monolingual children (bilingual: M = 0.71, SE = 0.02; monolingual: M = 0.67, SE = 0.02; Figure 2B).

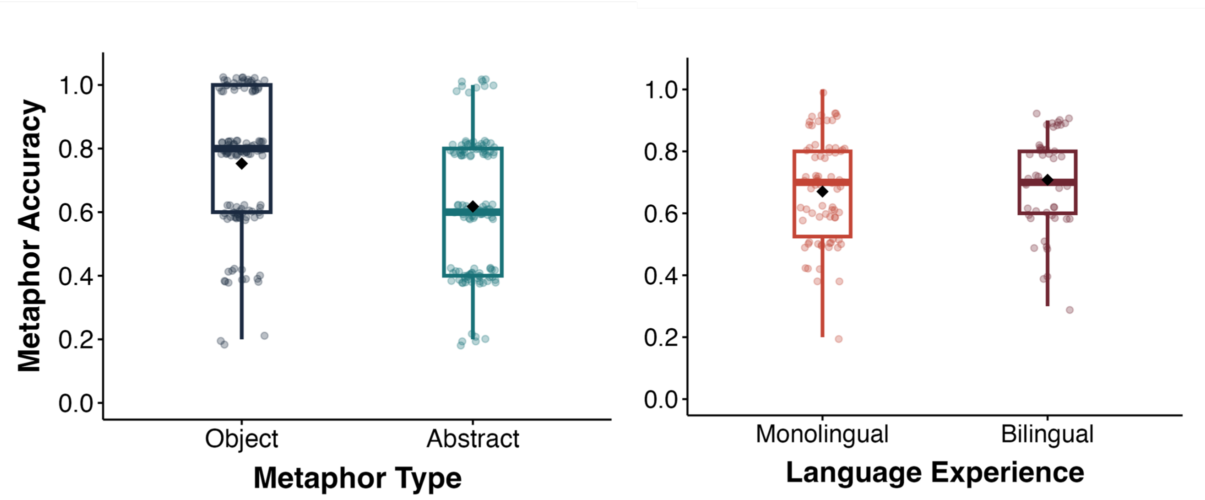

Figure 2. ***Experiment 1 metaphor comprehension accuracy by metaphor type (A) and language experience (B). Diamonds indicate group means.***

Vocabulary Knowledge

Because children needed to score above 70% on their understanding of the literal meaning of the words used in the metaphorical phrases to be included in the final analyses, children's vocabulary comprehension was very strong, and vocabulary questions were answered with an average accuracy of 93%. A linear model predicting vocabulary accuracy found that language learning environment was not a significant predictor of accuracy (B = 0.005, p = 0.74; monolingual children: M = 0.93, SE = 0.01; bilingual children: M = 0.92, SE = 0.01). However, participants scored significantly better on the vocabulary questions that corresponded to object metaphor questions compared to vocabulary questions that corresponded to abstract metaphor questions (B = 0.04, p = 0.007; object: M = 0.95, SE = 0.01; abstract: M = 0.91, SE = 0.01). As planned in our pre-registration, to determine how differences in vocabulary comprehension influence metaphor comprehension, we re-ran our original model predicting metaphor comprehension performance but analyzed only data from metaphor questions for which the corresponding vocabulary question was answered correctly. With this more stringent criterion, we found that there was a main effect of metaphor type (B = 0.62, p = .034), indicating higher performance on object metaphors compared to abstract metaphors, but the effect of language environment was still not significant (B = -0.31, p = .12).

Follow-up Analyses Investigating Age and Second Language Exposure

In a follow-up pre-registered analysis, we explored how metaphor comprehension was affected by age. When we included age as an additional continuous fixed effect, it was a significant positive predictor of performance (B = 0.42, p = .001), though again the effect of metaphor type was borderline (B = 0.49, p = .057), and the effect language learning environment was not significant (B = -0.30, p = .10). With age, children became increasingly successful in matching the metaphors to the correct image. In a second pre-registered follow-up model, we used percent exposure to a second language as a continuous predictor, rather than the dichotomous monolingual versus bilingual categorization. The outcome of this model again revealed a significant effect of age (B = 0.35, p = .007), as well as a significant effect of metaphor type (B = 0.70, p < .001). However, percent exposure to a second language was not a significant predictor (B = 0.11, p = .24). Note that this analysis involves a slightly smaller sample size (n = 112) due to some parents providing ranges of language exposure for their child rather than single values.

In addition to younger children having more difficulty with the metaphor questions, we also found that younger children were more likely to have their data excluded in comparison to older children. Of the 36 participants excluded for performance on vocabulary questions and attention check questions, 29 were under 3.5 years of age. Overall, 53% of all children under 3.5 were cut from further analyses due to poor performance on vocabulary and attention check questions. In a follow-up exploratory analysis, we ran a model using only data from children older than 3.5 years and found that metaphor type was a significant predictor of performance (B = 0.77, p = .009), but neither language learning environment (B = -0.34, p = .16) nor age (B = 0.45, p = .12) were significant predictors. Overall, these patterns of results indicate that children are able to comprehend both object and abstract metaphors, but additional data is needed to determine if abstract metaphors are more difficult. In addition, monolingual and bilingual children are equally proficient in comprehending metaphors.

Experiment 2

In Experiment 1, we tested whether objects metaphors are easier for preschool children to understand in comparison to abstract metaphors and whether children's language learning experience affects their metaphor comprehension. We found that children were successful in matching metaphorical expressions to the target picture for both object and abstract metaphors, but that abstract metaphorical phrases were potentially more difficult. We also found that children's proficiency with metaphor comprehension was not predicted by their language learning environment. In Experiment 2, we tested the robustness of these results with a conceptual replication conducted using Lookit (now called Children Helping Science; Scott & Schulz, 2017; Sheskin et al., 2020). Children Helping Science is a platform that enables families to participate in experiments over the web without an experimenter present. In addition to changing from a moderated to an unmoderated experiment format, we made a few other adjustments. First, we decided to focus on children aged 3.5 to 4.5 years (rather than 2.5 to 4.5 years) because these slightly older children are more likely to know the literal meaning of our metaphorical words and to stay engaged throughout the experimental session. In addition, we replaced the metaphorical expressions for which children performed at chance with novel metaphors and images to determine of the initial results were specific to the metaphorical phrases chosen. Finally, because children also performed better on the object vocabulary questions relative to the abstract vocabulary questions, we adopted a more stringent analysis strategy in which we only analyzed metaphor comprehension data for questions that children answered the corresponding vocabulary question correctly. This enabled us to confirm that differences in metaphor comprehension across metaphor types reflected metaphor comprehension itself versus comprehension of the literal meanings of the words.

Method

Participants

Data from 149 participants aged 3.5-4.5 years old (M = 3.89 years, SD = 0.32, 79 female) were included in the final analyses. As in Experiment 1, participants were classified as monolingual if they were exposed to English at least 90% of the time (n = 93) and classified as bilingual if they were exposed to a second language at least 20% of the time (n = 56). We preregistered a sample size of 66 monolingual children and 66 bilingual children. Sixty-six participants per language learning group gives us 90% power to detect a medium effect size for a paired samples *t*-test comparing metaphor accuracy between language experience groups (the effect size for a *t*-test comparing metaphor accuracy between the language experience groups in older children in Experiment 1 was $d = 0.57$). However, due to difficulties with recruiting bilingual participants, continued recruitment resulted in a larger sample of monolingual participants while still not being able to reach our target number of bilingual participants, which limits our statistical power. Caregivers of monolingual participants identified their children as 62% White, 27% multiracial, 10% Asian, and 1% Black or African American. With respect to annual household income, 33% of monolingual caregivers reported less than $100,000, 35% reported $100,000-$200,000, 25% reported more than $200,000, and 7% did not answer. Caregivers of bilingual participants identified their children as 36% Asian, 32% white, 23% multiracial, 5% Hispanic or Latine, and 4% chose not to respond. In terms of annual household income, 30% of bilingual caregivers reported less than 100k, 27% reported $100,000-$200,000, 30% reported more than $200,000, and 13% did not answer. The monolingual and bilingual language groups differed in terms of racial demographics ($X^2(3) = 23.73$, $p < .001$), but not in terms of household income ($X^2(2) = 1.21$, $p = .55$).

Data from an additional 36 participants were excluded for not fitting into our language criteria ($n = 13$), performing below 65% on the vocabulary questions ($n = 8$), not completing the study ($n = 7$), parent-reported developmental disabilities ($n = 3$), caregiver interference ($n = 3$), or performing below 66% on the attention checks ($n = 2$). The University of Washington Institutional Review Board approved of this study. Parents of the participants consented for their child's participation in the study and for video and audio recording of the study. Parents of participants were compensated with a $5 electronic Tango gift card. Data were collected between November 2021 and June 2023.

Procedure

This study was conducted entirely over Lookit (now called Children Helping Science; Scott & Schulz, 2017; Sheskin et al., 2020). Children completed the study with a parent, and there was no experimenter present. Parents read each question aloud to their child and clicked on the picture that their child chose. Because children could now point to their response, the Elmo and Cookie Monster images were removed from the stimuli. The procedure was otherwise identical to Experiment 1.

Materials and Design

The materials were identical to those used in Experiment 1, with the exception that five metaphor questions (one object and four abstract) in Experiment 1 for which children scored below 65% were replaced with new stimuli (Table 1). Because one of our main experimental questions is whether abstract metaphors are more difficult for children to understand than object metaphors, replicating our study with new metaphorical stimuli is critical for determining the robustness of our results. The mean ages of acquisition (Kuperman et al., 2012) for the target object and abstract metaphor words were not statistically different ($t(10.5) = 0.70$, $p = .50$). The new metaphor stimuli were validated in a sample of English-speaking adults (N = 30) who chose the target image 100% of the time. However, initial analyses revealed that children's performance on one version of a new metaphor question ("which wave is weaker?") was significantly below chance, and this item (both the A and B versions) was therefore excluded from all subsequent analyses. The stimuli used in Experiment 2, along with children's accuracy on the metaphor and vocabulary version of the questions, are presented in Table 2.

Table 2
Mean Metaphor and Vocabulary Comprehension Accuracy by Question

Question "Which picture has the…"	**Version**	**Question Type**	**Mean Metaphor Accuracy [95% CI]**	**Mean Vocab Accuracy [95% CI]**
Bottle with a big belly		Object	0.88 [0.75-1]	0.87 [0.75-0.99]
Cup with a sweater		Object	0.80 [0.69-0.91]	0.89 [0.77-1]
Dog with socks		Object	0.91 [0.78-1]	0.95 [0.82-1]
Horse with a back-pack		Object	0.81 [0.70-0.92]	0.97 [0.84-1]
Tree with hair		Object	0.78 [0.67-0.89]	0.97 [0.84-1]
Girl who is feeling up	A	Abstract	0.86 [0.69-1]	0.96 [0.77-1]
Girl who is feeling down	B	Abstract	0.66 [0.54-0.78]	1 [0.82-1]
Phone that is healthy	A	Abstract	0.63 [0.50-0.76]	0.99 [0.79-1]
Phone that is sick	B	Abstract	0.87 [0.69-1]	0.75 [0.61-0.89]
Wave that is stronger*	A	Abstract	0.97 [0.77-1]	0.94 [0.75-1]
Wave that is weaker*	B	Abstract	0.36 [0.27-0.45]	0.52 [0.42-0.62]
Boy who is having a bumpy day	A	Abstract	0.85 [0.67-1]	0.90 [0.72-1]
Boy who is having a smooth day	B	Abstract	0.68 [0.54-0.82]	0.75 [0.61-0.89]
Flower that is happy	A	Abstract	0.96 [0.77-1]	0.97 [0.78-1]
Flower that is sad	B	Abstract	0.93 [0.75-1]	0.95 [0.78-1]

Note. *Asterisks denote items with mean metaphor comprehension accuracy below 50% for* *version B. This question (both the A and B versions) was dropped from all subsequent analyses.*

Data Analysis

The preregistration for this study can be found at https://osf.io/ta8vf. Because we only planned to analyze metaphor questions for which children answered the corresponding vocabulary question correctly, we used a slightly more lenient vocabulary performance threshold of at least 65% correct (rather than 75% in Experiment 1). On average, children contributed 8.3 metaphor questions to the analyses. As in Experiment 1, our final statistical models deviated slightly from the preregistered models, and we used models with the maximal random effects structure that did converge. Data cleaning, analyses, and visualizations were performed in R using the tidyverse and lmerTest packages (Kuznetsova et al., 2017; Wickham et al., 2019).

Results

Main Analyses

As in Experiment 1, we ran a pre-registered logistic mixed effects model predicting trial accuracy with metaphor type (object or abstract), language learning environment (monolingual or bilingual), the interaction between metaphor type and language environment, and question version (A or B) as fixed effects, and random intercepts for subject and item. This model revealed no significant main effects or interactions (metaphor: $B = 0.35$, $p = .43$; language environment: $B = 0.21$, $p = .43$; interaction: $B = -0.38$, $p = .24$). Children matched the metaphorical phrase to the correct picture at above chance levels for both types of metaphors, as they did in Experiment 1 ($ts > 16.28$, $ps < 0.001$). In contrast to the results of Experiment 1, however, performance on object metaphor questions was equivalent to performance on abstract metaphor questions (object: $M = 0.83$, $SE = 0.02$; abstract: $M = 0.80$, $SE = 0.02$; Figure 3A). In addition, bilingual children performed equivalently to monolingual children (bilingual: $M = 0.82$, $SE = 0.02$; monolingual: $M = 0.82$, $SE = 0.01$; Figure 3B).

In a follow-up model, we added age as an additional predictor and found that metaphor performance improved with age ($B = 0.94$, $p = .003$), but no other main effects or interactions were significant ($Bs < 0.39$, $ps > .23$). In a pre-registered follow-up model, we used percent exposure to a second language as a continuous predictor, rather than the dichotomous monolingual versus bilingual categories. This model used a slightly smaller sample size due to some parents providing ranges of second language exposure (n = 139). The outcome of this model again revealed only a significant effect of age ($B = 0.95$, $p = .006$), and percent exposure to a second language was not a significant predictor ($B = -0.04$, $p = .75$). These patterns of results suggest that children are equally proficient at understanding both object and abstract metaphors and metaphor comprehension improves with age, but language learning environment does not influence metaphor comprehension.

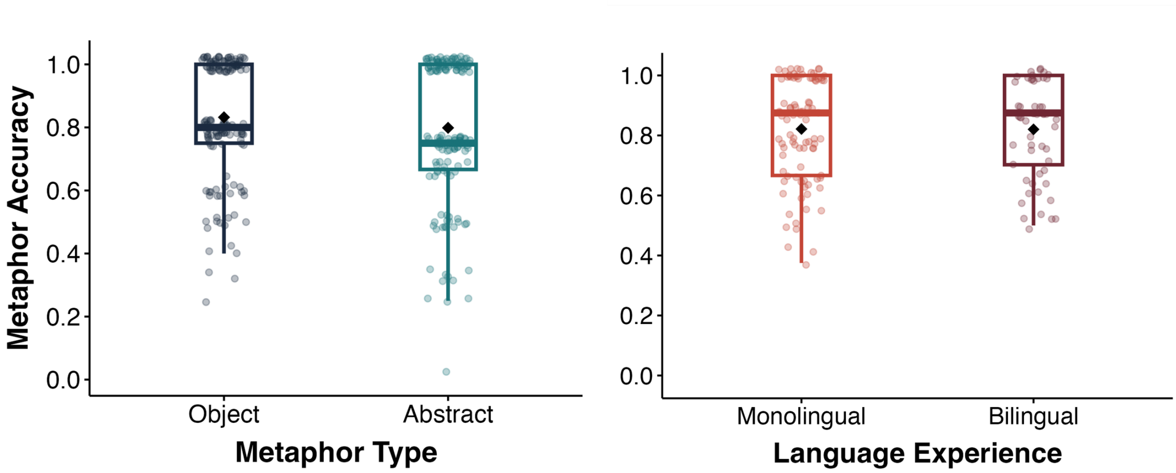

Figure 3. ***Experiment 2 metaphor comprehension accuracy by metaphor type (A) and language experience (B). Diamonds indicate group means.***

Vocabulary Knowledge

Children's vocabulary comprehension performance was slightly better for object metaphor stimuli compared to abstract metaphor stimuli ($B = 0.382$, $p = .068$; object: $M = 0.93$, $SE = 0.01$; abstract: $M = 0.91$, $SE = 0.01$). As a result, slightly more abstract than object metaphor questions were dropped due to missing the corresponding vocabulary questions in the original analysis for the two types of words (56 abstract questions versus 49 object questions). Vocabulary comprehension did not differ between monolingual and bilingual children ($B = 0.32$, $p = .191$; monolingual: $M = 0.93$, $SE = 0.01$; bilingual: $M = 0.91$, $SE = 0.01$).

Combined Analyses Using Data from Experiments 1 and 2

In a final exploratory analysis, we combined the data from Experiments 1 and 2 to see if increasing the sample size might reveal an effect of language learning environment. We used a logistic mixed effects model predicting trial accuracy with metaphor type, language learning environment, age, their interactions, and experiment as fixed effects, along with random intercepts for subject and item. For Experiment 1 data, we used data only from metaphor comprehension questions for which children answered the corresponding vocabulary question correctly and only from children aged 3.5 years and older to make it comparable to Experiment 2 data. The outcome of this model revealed a significant of age (B = 0.84, p < .001), indicating the metaphor performance improved with age. The effect of experiment was also significant (B = 0.58, p < .001), indicating the performance was higher in Experiment 2 compared to Experiment 1. No other main effects or interactions were significant (Bs < 0.50, ps > .13). Taken together, the results of Experiments 1 and 2 suggest that children are equally proficient at understanding object and abstract metaphors, but language learning environment does not influence metaphor comprehension.

General Discussion

The current study was designed to increase our understanding of metaphorical language comprehension in preschoolers in two key ways. First, given preschooler's success in comprehending metaphors based on shared perceptual and functional features of objects (Pouscoulous & Tomasello, 2020; Zhu et al., 2024; Zhu & Gopnik, 2023), we tested whether preschool children are also able to understand metaphors that do not rely on object knowledge. We compared children's performance on metaphors based on shared features of objects, which we term object metaphors, with their performance on abstract metaphors that involve extending words across modalities and conceptual domains. We predicted that children would be able to understand both types of metaphors, but that they would be more successful in understanding object metaphors compared to abstract metaphors. The second goal of the study was to explore how children's language learning environments affect metaphor comprehension, because the majority of work on children's metaphor comprehension has focused only on monolingual children. We predicted that because of their regular experience with mapping more than one label to the same concept, bilingual children might be more successful in understanding metaphorical language relative to monolingual children. We addressed these questions in a primary experiment (Experiment 1) and a conceptual replication (Experiment 2) to increase the strength of our conclusions.

Prior studies have shown that preschool-aged children can understand metaphors that involve mapping between shared perceptual and functional features – in other words, metaphors in which children can employ their knowledge of objects (Gentner, 1988; Pouscoulous & Tomasello, 2020; Vosniadou et al., 1984; Zhu et al., 2024; Zhu & Gopnik, 2023). On the other hand, when faced with abstract metaphors, such as those involving metaphorical motion (e.g., time *flies*), younger preschoolers are less likely to succeed, and comprehension increases between the ages of 3 and 5 years (Özçalişkan, 2005; Özçalışkan, 2007, Sites & Özçalışkan, 2012). These findings indirectly suggest that the ability to understand object metaphors may develop before the ability to understand abstract metaphors.

In the present study, we tested preschool children's comprehension of object and abstract metaphors. Across both experiments, we found that preschoolers were successful in comprehending both types of metaphors. Although in Experiment 1 we found some evidence that children might be more accurate in comprehending object metaphors, in Experiment 2, children were equally accurate in comprehending both types of metaphors. In Experiment 1, there were several metaphorical phrases for which children's comprehension accuracy was under 65%. In Experiment 2, we replaced these phrases. These new metaphorical phrases were apparently easier for children to comprehend, as evidenced by higher performance, which suggests that the disparate results across the two experiments is due to the specific stimuli used. Consistent with prior work (P. Bloom, 2000; Braginsky et al., 2019; Gentner & Boroditsky, 2001), we found that children's comprehension of the literal meaning of the target words was higher for the object metaphor words, which were concrete nouns, relative to their comprehension of the literal meaning of the target abstract metaphor words, which were adjectives. However, when children understood the literal meaning of the target words, they were able to understand metaphors that evoke similarities beyond shared properties and functions of objects. Taken together, the results of Experiments 1 and 2 suggest that preschool-aged children can comprehend not only metaphors based on shared object features and functions, but also metaphors based on more abstract relations.

To resolve object metaphors, children can use their knowledge about the objects involved. For example, when hearing the phrase, "the dog with the white socks," children can use their knowledge that socks go on feet to cue their attention to the dogs' feet in the two image choices. For abstract metaphors, there is no direct physical correlate that supports the metaphor. For example, when hearing the phrase, "the girl is feeling down," the word *down* is unlikely to automatically cue children's attention towards emotions. However, this association between vertical space and emotional valence is common in many cultures (Gibbs, 1994; Lakoff & Johnson, 1980). More broadly, there is a large body of work demonstrating that infants and young children are sensitive to a variety of cross-modal and cross-domain associations (e.g., Casasanto & Henetz, 2012; de Hevia et al., 2014; Dolscheid et al., 2014; Lourenco & Longo, 2010; Srinivasan & Carey, 2010; Starr & Srinivasan, 2018; Walker et al., 2010). Children's intuitions about these associations, for example associating a loud sound with visual brightness, likely guide their abstract metaphor comprehension performance. In addition, as demonstrated by the example "the girl is feeling *down*," many abstract metaphors involve a mapping between space and another domain. In English, for example, we use spatial proximity to describe similarity and personal relationships (e.g., "the two colors were a *close* match," "she is a *distant* relative") and use spatial positions and distances to represent time (e.g., "I'm looking *forward* to a vacation," "the meeting was *short*"). This pattern of using spatial words to describe more abstract phenomena is common across languages and cultures (Clark, 1973; Lakoff & Johnson, 1980; Sweetser, 1990; Talmy, 1988; Xu et al., 2017). With respect to vocabulary development, this pattern of metaphorically extending already-known words for concrete concepts (e.g., spatial words) to more abstract concepts (e.g., time) may help children solve the otherwise difficult mapping problem for learning labels for abstract concepts (Starr et al., 2021). Therefore, abstract metaphors might not only be understandable by young children, they might also help children learn about new concepts, similar to how preschool-aged children can learn from novel metaphors involving objects (Zhu & Gopnik, 2023).

Our abstract metaphors were relatively heterogenous and involved a variety of topics including cross-modal mappings, emotions, and the tenor of a day, and some of the metaphors were easier for children than others. Not surprisingly, previous studies have also found item-level variation in children's metaphor comprehension across different metaphorical phrases of the same type and structure (e.g., Pouscoulous & Tomasello, 2020; Winner et al., 1980). We aimed to choose metaphors that involved concepts that would be familiar to preschool-aged children, but it is clear that some phrases were more difficult than others. An important direction for future work will therefore be to investigate different types of abstract metaphors. In particular, metaphors involving cross-modal mappings may be easier for children to understand than metaphors involving psychological targets (Winner et al., 1976). Because understanding metaphors requires making inferences about a speaker's intent, it has been argued that understanding metaphors requires theory of mind (Happé, 1993). However, more recent work that has suggested that not all metaphors equally involve theory of mind, and theory of mind might be particularly relevant for understanding metaphors involving mental states (Lecce et al., 2019) and less involved in understanding metaphors based on object properties (Norbury, 2005). Additional work using a broader range of metaphorical types is therefore needed to better understand factors beyond metaphor type that contribute to the relative difficulty of understanding metaphorical language.

A second focus of the present study was to investigate how children's early language learning environments influence metaphor comprehension. Although rates of bilingualism are increasing around the world, most studies of metaphor comprehension have focused on monolingual children. We hypothesized that because bilingual children are continuously learning that concepts can have more than one label, they may have an advantage in metaphor comprehension in comparison to their monolingual peers. However, our data do not support this hypothesis: monolingual and bilingual children were equally proficient at understanding both object and abstract metaphors. Although each of our experiments on their own might be underpowered to detect a small effect of language learning experience, we also found no significant effect when we collapsed across the two data sets to double the sample size. Critically, in neither experiment did bilingual children perform worse than their monolingual peers. This pattern of results suggests that even though bilingual preschoolers know fewer English words than their monolingual peers (Hoff et al., 2012), they are able to use the words they do know in complex ways.

For our primary analyses, we characterized children's language status as monolingual or bilingual and used parent report to decide to which category a child was assigned. Although this binary categorization is common in language development studies, it flattens the true experience of bilingualism (Byers-Heinlein et al., 2019; Rocha-Hidalgo & Barr, 2022). In follow-up analyses, we used children's percent exposure to a second language as a continuous predictor rather than using the categorical monolingual versus bilingual distinction. As with the models using the categorical classification, we again did not find that language learning environment significantly predicted metaphor comprehension. However, this continuous measure still fails to capture many meaningful differences in children's language learning environments, such as whether a child is learning two languages simultaneously or sequentially or what the sources of exposure are to each language. We also did not measure children's proficiency in either language beyond children's comprehension of target words in English. It is possible that with more nuanced measures of a child's language learning environment and proficiency, significant effects may be found or that effects may emerge with age as children's linguistic abilities develop further and children encounter other forms of metaphors. In addition, because children in these studies came from primarily middle- and upper-class families in the US, caution must be used in generalizing these findings to children growing up in other contexts.

The current work used a paradigm adopted from Pouscoulous and Tomasello (2020), in which children were asked to choose which of two pictures best matched a given metaphorical phrase. The paradigms used by Zhu and colleagues (2023, 2024) also involved visual depictions of metaphorical language. The use of visuals likely supports early metaphor comprehension and may contribute to the differences in metaphorical understanding found in these studies compared to those conducted by Özçalışkan (2005, 2007), which involved verbal descriptions only. An important direction for future work will be to investigate the contexts that facilitate versus hinder preschooler's metaphorical language comprehension. In addition, we used a two-alternative forced choice paradigm in which the distractor image was either an unrelated object feature (for object metaphors) or the opposite conceptual mapping (for abstract metaphors). To further understand children's early metaphor comprehension, researchers may consider asking children to choose between images depicting metaphorical and literal interpretations or ask them to explain their answers.

Metaphors are a powerful linguistic tool – they enable speakers to highlight similarities and make comparisons by flexibly reusing known words in novel contexts. Although metaphorical language can be difficult for young children, recent work has demonstrated that preschool-aged children can understand and even learn from novel metaphors (Pouscoulous & Tomasello, 2020; Zhu et al., 2024; Zhu & Gopnik, 2023). Here, we found that 3- and 4-year-old children can comprehend not just perceptual and functional metaphors involving objects, but also abstract metaphors that involve less tangible domains. These results contribute to a growing body of work demonstrating that preschool children are able to understand multiple types of non-literal language, including metaphors, metonyms, and irony (Di Paola et al., 2020; Falkum, 2022; Pouscoulous, 2023; Pouscoulous & Tomasello, 2020; Starr et al., 2021; Zhu et al., 2024, 2024). During the preschool years, as children are rapidly learning new words, they are also able to use these words in flexible and sophisticated ways.

References

Andrews, M. L., & Madeira, S. S. (1977). The Assessment of Pitch Discrimination Ability in Young Children. *Journal of Speech and Hearing Disorders, 42*(2), 279–286. https://doi.org/10.1044/jshd.4202.279

Ben-Zeev, S. (1977). The Influence of Bilingualism on Cognitive Strategy and Cognitive Development. *Child Development, 48*(3), 1009–1018. https://doi.org/10.2307/1128353

Bloom, L., Lightbown, P., Hood, L., Bowerman, M., Maratsos, M., & Maratsos, M. P. (1975). Structure and variation in child language. *Monographs of the Society for Research in Child Development, 40*(2), 1. https://doi.org/10.2307/1165986

Bloom, P. (2000). *How children learn the meanings of words*. MIT Press.

Bowdle, B. F., & Gentner, D. (2005). The Career of Metaphor. *Psychological Review, 112*(1), 193–216. https://doi.org/10.1037/0033-295X.112.1.193

Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. C. (2019). Consistency and Variability in Children's Word Learning Across Languages. *OPEN MIND,* 16. https://doi.org/10.1162/opmi_a_00026

Byers-Heinlein, K. (2017). Bilingualism affects 9-month-old infants' expectations about how words refer to kinds. *Developmental Science, 20*(1), e12486. https://doi.org/10.1111/desc.12486

Byers-Heinlein, K., Esposito, A. G., Winsler, A., Marian, V., Castro, D. C., & Luk, G. (2019). The Case for Measuring and Reporting Bilingualism in Developmental Research. *Collabra: Psychology, 5*(1), 37. https://doi.org/10/gg6xjh

Casasanto, D., Fotakopoulou, O., & Boroditsky, L. (2010). Space and time in the child's mind: Evidence for a cross-dimensional asymmetry. *Cognitive Science, 34*(3), 387–405. https://doi.org/10.1111/j.1551-6709.2010.01094.x

Casasanto, D., & Henetz, T. (2012). Handedness Shapes Children's Abstract

Concepts. *Cognitive Science, 36*(2), 359–372. https://doi.org/10/bp9jg4

Clark, H. H. (1973). Space, time, semantics, and the child. In *Cognitive development and the acquisition of language* (pp. 27–63). Academic Press.

de Hevia, M. D., Izard, V., Coubart, A., Spelke, E. S., & Streri, A. (2014). Representations of space, time, and number in neonates. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1323628111

Di Paola, S., Domaneschi, F., & Pouscoulous, N. (2020). Metaphorical developing minds: The role of multiple factors in the development of metaphor comprehension. *Journal of Pragmatics, 156*, 235–251. https://doi.org/10/gg7p8d

Diaz, V., & Farrar, M. J. (2018). The missing explanation of the false-belief advantage in bilingual children: A longitudinal study. *Developmental Science, 21*(4), e12594. https://doi.org/10.1111/desc.12594

Dolscheid, S., Hunnius, S., Casasanto, D., & Majid, A. (2014). Prelinguistic Infants Are Sensitive to Space-Pitch Associations Found Across Cultures. *Psychological Science, 25*(6), 1256–1261. https://doi.org/10.1177/0956797614528521

Falkum, I. L. (2022). The development of non-literal uses of language: Sense conventions and pragmatic competence. *Journal of Pragmatics, 188*, 97–107. https://doi.org/10.1016/j.pragma.2021.12.002

Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Monographs of the Society for Research in Child Development, 49*(4), 988. https://doi.org/10.2307/1128738

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child Development, 59*, 47–59. https://doi.org/10.2307/1130388

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). Cambridge University Press. https://doi.org/10.1017/CBO9780511620669.010

Gibbs, R. W. (1994). *The Poetics of Mind*. Cambridge University Press.

Groba, A., De Houwer, A., Obrig, H., & Rossi, S. (2019). Bilingual and Monolingual First Language Acquisition Experience Differentially Shapes Children's Property Term Learning: Evidence from Behavioral and Neurophysiological Measures. *Brain Sciences, 9*(2), 40. https://doi.org/10/ghv59w

Halberda, J. (2003). The development of a word-learning strategy. *Cognition, 87*(1), B23–B34. https://doi.org/10/dbxw3d

Happé, F. G. E. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition, 48*(2), 101–119. https://doi.org/10.1016/0010-0277(93)90026-R

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language, 39*(1), 1–27. https://doi.org/10.1017/S0305000910000759

Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language Experience Shapes the Development of the Mutual Exclusivity Bias. *Infancy, 15*(2), 125–150. https://doi.org/10/d42m2d

Jardak, A., & Byers-Heinlein, K. (2019). Labels or Concepts? The Development of Semantic Networks in Bilingual Two-Year-Olds. *Child Development, 90*(2), e212–e229. https://doi.org/10/gjkzwt

Kuhn, T. S. (1993). Metaphor in science. In A. Ortony, *Metaphor and thought* (pp. 533–542). Cambridge University Press. https://doi.org/10.1017/CBO9781139173865.024

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. The University of Chicago Press.

Lecce, S., Ronchi, L., Del Sette, P., Bischetti, L., & Bambini, V. (2019). Interpreting physical and mental metaphors: Is Theory of Mind associated with pragmatics in middle childhood? *Journal of Child Language, 46*(2), 393–407. https://doi.org/10.1017/S030500091800048X

Lourenco, S. F., & Longo, M. R. (2010). General magnitude representation in human infants. *Psychological Science, 21*(6), 873–881. https://doi.org/10.1177/0956797610370158

Luk, G. (2017). Bilingualism. In B. Hopkins, E. Geangu, & S. Linkenauger (Eds.), *The Cambridge Encyclopedia of Child Development* (2nd ed., pp. 385–391). Cambridge University Press. https://doi.org/10.1017/9781316216491.062
Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20*(2), 121–157. https://doi.org/10.1016/0010-0285(88)90017-5

Norbury, C. Frazier. (2005). The relationship between theory of mind and metaphor: Evidence from children with language impairment and autistic spectrum disorder. *British Journal of Developmental Psychology, 23*(3), 383–399. https://doi.org/10.1348/026151005X26732

Özçalişkan, Ş. (2005). On learning to draw the distinction between physical and metaphorical motion: Is metaphor an early emerging cognitive and linguistic capacity? *Journal of Child Language, 32*(2), 291–318. https://doi.org/10.1017/S0305000905006884

Özçalışkan, Ş. (2007). Metaphors We Move By: Children's Developing Understanding of Metaphorical Motion in Typologically Distinct Languages. *Metaphor and Symbol, 22*(2), 147–168. https://doi.org/10.1080/10926480701235429

Pouscoulous, N. (2011). Metaphor: For adults only? *Belgian Journal of Linguistics, 25,* 51–79. https://doi.org/10.1075/bjl.25.04pou

Pouscoulous, N. (2023). More than one path to pragmatics? Insights from children's grasp of implicit, figurative and ironical meaning. *Cognition, 240,* 105531. https://doi.org/10.1016/j.cognition.2023.105531

Pouscoulous, N., & Tomasello, M. (2020). Early birds: Metaphor understanding in 3-year-olds. *Journal of Pragmatics, 156,* 160–167. https://doi.org/10.1016/j.pragma.2019.05.021

Repnik, K. M., Chondrogianni, V., & Sorace, A. (2021). Linking disambiguation and retention in a developmental eye-tracking study with monolingual and multilingual children. *Journal of Experimental Child Psychology, 206,* 105072. https://doi.org/10.1016/j.jecp.2020.105072

Rocha-Hidalgo, J., & Barr, R. (2022). Defining bilingualism in infancy and toddlerhood: A scoping review. *International Journal of Bilingualism,* 13670069211069067. https://doi.org/10.1177/13670069211069067

Rubio-Fernández, P., & Grassmann, S. (2016). Metaphors as Second Labels: Difficult for Preschool Children? *Journal of Psycholinguistic Research, 45*(4), 931–944.

https://doi.org/10.1007/s10936-015-9386-y

Scott, K., & Schulz, L. (2017). Lookit (Part 1): A New Online Platform for Developmental Research. *Open Mind, 1*(1), 4–14. https://doi.org/10/gg84jq

Sheskin, M., Scott, K., Mills, C. M., Bergelson, E., Bonawitz, E., Spelke, E. S., Fei-Fei, L., Keil, F. C., Gweon, H., Tenenbaum, J. B., Jara-Ettinger, J., Adolph, K. E., Rhodes, M., Frank, M. C., Mehr, S. A., & Schulz, L. (2020). Online Developmental Science to Foster Innovation, Access, and Impact. *Trends in Cognitive Sciences,* S1364661320301455. https://doi.org/10.1016/j.tics.2020.06.004

Silberstein, L., Gardner, H., Phelps, E., & Winner, E. (1982). Autumn leaves and old photographs: The development of metaphor preferences. *Journal of Experimental Child Psychology, 34*(1), 135–150. https://doi.org/10.1016/0022-0965(82)90036-4

Srinivasan, M., & Carey, S. (2010). The long and the short of it: On the nature and origin of functional overlap between representations of space and time. *Cognition, 116*(2), 217–241. https://doi.org/10.1016/j.cognition.2010.05.005

Starr, A., Cirolia, A. J., Tillman, K. A., & Srinivasan, M. (2021). Spatial Metaphor Facilitates Word Learning. *Child Development, 92*(3), e329–e342. https://doi.org/10.1111/cdev.13477

Starr, A., & Srinivasan, M. (2018). Spatial Metaphor and the Development of Cross-Domain Mappings in Early Childhood. *Developmental Psychology, 54*(10), 1822–1832. https://doi.org/10.1037/dev0000573

Stites, L. J., & Özçalışkan, Ş. (2012). Developmental changes in children's comprehension and explanation of spatial metaphors for time. *Journal of Child Language, 40*(05), 1123–1137. https://doi.org/10.1017/S0305000912000384

Sweetser, E. (1990). *From Etymology to Pragmatics*. Cambridge University Press. https://doi.org/10.1017/CBO9780511620904

Talmy, L. (1988). Force Dynamics in Language and Cognition. *Cognitive Science, 12*(1), 49–100. https://doi.org/10.1207/s15516709cog1201_2

Vosniadou, S. (1987). Children and Metaphors. *Child Development, 58*(3), 870–885. https://doi.org/10.2307/1130223

Vosniadou, S., Ortony, A., Reynolds, R. E., & Wilson, P. T. (1984). Sources of difficulty in the young child's understanding of metaphorical language. *Monographs of the Society for Research in Child Development, 55*(4), 1588.

https://doi.org/10.2307/1130028

Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal Infants' Sensitivity to Synaesthetic Cross-Modality Correspondences. *Psychological Science*, *21*(1), 21–25. https://doi.org/10.1177/0956797609354734

Wickham, H. (2014). Tidy data. *Journal of Statistical Software*.

Winner, E., Engel, M., & Gardner, H. (1980). Misunderstanding metaphor: What's the problem? *Journal of Experimental Child Psychology, 30*(1), 22–32. https://doi.org/10.1016/0022-0965(80)90072-7

Winner, E., McCarthy, M., Kleinman, S., & Gardner, H. (1979). First metaphors. *New Directions for Child and Adolescent Development*, *1979*(3), 29–41. https://doi.org/10.1002/cd.23219790305

Winner, E., Rosenstiel, A. K., & Gardner, H. (1976). The development of metaphoric understanding. *Developmental Psychology, 12*(4), 289–297. https://doi.org/10.1037/0012-1649.12.4.289

Xu, Y., Malt, B. C., & Srinivasan, M. (2017). Evolution of word meanings through metaphorical mapping: Systematicity over the past millennium. *Cognitive Psychology, 96*, 41–53. https://doi.org/10.1016/j.cogpsych.2017.05.005

Zhu, R., Goddu, M. K., Zhu, L. Z., & Gopnik, A. (2024). Preschoolers' Comprehension of Functional Metaphors. *Open Mind: Discoveries in Cognitive Science*, *8*, 924–949. https://doi.org/10.1162/opmi_a_00152

Zhu, R., & Gopnik, A. (2023). Preschoolers and Adults Learn From Novel Metaphors. *Psychological Science, 34*(6), 696–704. https://doi.org/10.1177/09567976231165267

Data, code and materials availability statement

The data and code necessary to reproduce the analyses presented here are publicly accessible, as are the materials necessary to attempt to replicate the findings. Analyses were also pre-registered. Data, code, materials, and the preregistration for this research are available at the following URL: https://osf.io/ta8vf.

Ethics statement

Ethics approval was obtained from the institutional review board at the University of Washington. Parents of all participants provided informed consent.

Authorship and Contributorship Statement

AS conceptualized the study and wrote the first draft of the manuscript, AS and TP designed the materials, and TP and SS collected the data and contributed to revising the manuscript. All authors contributed to analyzing the data and approved the final version of the manuscript.

Acknowledgements

This research was supported by an award from the NSF Division of Research on Learning (award number 2201033) to AS. Thank you to the participating families and to Wendy Garcia, Elizabeth Li, Sam Seaver, Eun-Sun Shin, Julia Souza, and Yongyan Yue for their assistance with data collection.

License