Journal of Experimental Child Psychology 266 (2026) 106487

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

The role of epistemic reasoning in mutual exclusivity inferences

Khuyen N. Le *, David Barner

Department of Psychology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0109, United States

ARTICLE INFO

Keywords:
Mutual exclusivity
Word learning
Epistemic reasoning
Theory of Mind
Conventionality

ABSTRACT

When encountering a novel word, adults and children as young as 12 months old often reason that it refers to a novel object rather than one with an existing name – making a 'mutual exclusivity inference.' We explored the mechanism of this inference, aiming to differentiate between three hypotheses: whether mutual exclusivity arises due to reasoning about a specific speaker's knowledge, projection of one's own egocentric knowledge, or reasoning generally about the conventionality of labels. Adults and 3.5–5-year-old children in our experiment heard a label being either taught or invented on the spot. They were then asked to decide what another label, produced by a speaker who was absent during the first label's introduction, referred to. Our results revealed that adults and older children made more mutual exclusivity inferences when the first label was taught compared to when it was invented. Additionally, they were more likely to exclude the previously-labeled object in the referential choice if they judged that the speaker also knew the label, regardless of how the label was introduced. Both adults and older children also showed sensitivity to the speaker's explicit statements of knowledge, excluding an object that the speaker explicitly stated that he did not know the name of from their referential choice. These results suggested that adults and older children reasoned about a specific speaker's knowledge of labels in order to make mutual exclusivity inferences, likely in the form of a Gricean inference where they reasoned about possible alternative utterances a speaker could have said.

Introduction

When children learn words, they are faced with the problem of identifying the speaker's intended meaning from a large space of possible alternatives. How do children narrow their hypotheses to acquire words so quickly in early childhood? On one hypothesis, children limit their hypotheses via exclusion inferences, either by assuming that words contrast in meaning (Clark, 1988, 1990), or by assuming that mappings between words and their referents are mutually exclusive (Markman & Wachtel, 1988). For example, if presented with an object with a known label, like an apple, and a novel object like a papaya, a child might assume that a novel label, like 'papaya' maps onto the novel object, rather than onto the object that has an existing label. Consistent with this, previous studies find evidence of exclusion inferences in children as young as 12 months old (Diesendruck & Markson, 2001; Markman et al., 2003; Golinkoff et al., 1992; Pomiechowska et al., 2021; Liittschwager & Markman, 1994; Merriman & Stevenson, 1997; Merriman & Bowman, 1989; Halberda, 2003).

While it is well-accepted that children make exclusion inferences when learning novel words, there is an ongoing debate about the nature of these inferences. Inspired by Gricean maxims (e.g., Grice, 1975), some have argued that children make these inferences by

* Corresponding author.
E-mail address: khuyenle@ucsd.edu (K.N. Le).

https://doi.org/10.1016/j.jecp.2026.106487
Received 17 June 2025; Received in revised form 19 December 2025;
Available online 25 February 2026

reasoning about the speaker's epistemic states ('epistemic' accounts; Bloom, 2000; Clark, 1990; Diesendruck & Markson, 2001; Tomasello, 2000). According to these accounts, a child who hears an adult speaker say the word "papaya" in the presence of an apple and a novel fruit reasons that it would be uncooperative for the speaker to request the apple by saying 'papaya', because the speaker knows an alternative label for the apple. Therefore, the child infers that the speaker must intend to refer to the novel fruit with the novel label, as in (1) – (5) (see Clark, 1990; Gathercole, 1989).

(1) The speaker utters 'papaya' to request one of two fruits, one of which is an apple.
(2) I assume that person is cooperative, and is trying to communicate with me.
(3) I assume that the speaker is knowledgeable, and knows what apples are called.
(4) I infer that if the speaker had intended to request the apple, they would have uttered 'apple".
(5) I conclude that the speaker did not intend to request the apple, and therefore that 'papaya' must refer to the other object.

Compatible with this account, some studies have reported evidence that children are sensitive to speakers' intentional states when making exclusion inferences. For example, Diesendruck et al. (2010) found that children do not make mutual exclusivity inferences when speakers are shown to be ignorant of common labels (e.g., calling a cup a 'ball'). Similarly, mutual exclusivity inferences appear to be suspended for speakers of different languages. For example, in a study by Au and Glusman (1990), children who were taught a novel label by one experimenter then excluded that object when a second experimenter (absent during teaching) made a request using a different label. However, children chose randomly between the two objects when the two experimenters spoke different languages (e.g., English and Spanish; see also Diesendruck, 2005). On epistemic accounts, this result is predicted if children believe that speakers of different languages have different beliefs about words and their meanings.

In contrast to epistemic accounts, non-epistemic accounts argue that although children may ultimately learn to consider speaker states when learning words, they initially take a non-epistemic stance, and rely solely on their own knowledge of words to make exclusion inferences (Markman, 1990; Markman & Wachtel, 1988; see also Golinkoff et al., 1992; Merriman & Bowman, 1989). For example, children might assume that each word has a distinct meaning, and that novel words should not denote referents that have a known label, and should therefore refer to novel objects. According to this hypothesis, the child's reasoning in the 'papaya' scenario described above would proceed as in (6)–(9):

(6) The speaker utters 'papaya' to request one of two fruits, one of which is an apple.
(7) I assume that 'papaya' and 'apple' refer to different phenomena.
(8) I know that apples already have a name ('apple').
(9) Therefore, 'papaya' refers to objects that are not apples, i.e., the other object.

In support of non-epistemic accounts, there is evidence that children compute mutual exclusivity inferences by as early as 12 months of age (Pomiechowska et al., 2021), when theory of mind abilities are still rudimentary (see also Halberda, 2003; Markman et al., 2003; Mather & Plunkett, 2011; Houston-Price et al., 2010).[1] On such accounts, children may suspend mutual exclusivity inferences for speakers with incorrect beliefs or who speak a different language not because they are reasoning in a Gricean way, but because they have a more general heuristic to ignore or discount evidence from unreliable individuals (Koenig et al., 2004; Koenig & Harris, 2007; Sabbagh & Shafman, 2009; Sobel & Kushnir, 2013; Scofield & Behrend, 2008).

Conventionality assumptions in mutual exclusivity inferences

Critical to differentiating Gricean reasoning from simpler heuristics are methods that manipulate the speaker's knowledge of specific words, without characterizing them as globally ignorant or error-prone. Specifically, in the Gricean algorithm, the inference that 'papaya' must refer to a novel object hinges on the assumption that the speaker is knowledgeable of the alternative word 'apple'. Given this, the inference should be made for most speakers, but not for a speaker who has never heard the word 'apple'. In a clever test of this idea, Diesendruck and Markson (2001) introduced children to two objects and a puppet named Percy. After this initial introduction, Percy either stayed outside and listened to the conversation, or disappeared inside his house where he could not see or hear anything that happened. The experimenter then labeled one of the objects with either a novel word (e.g. 'dax') or a fact ("My cat likes to play with this."). After the labeling, Percy then reappeared if he wasn't present the entire time, and asked the child to give him an object with a different name or fact corresponding to the type of introduction used (e.g., "Can you give me the *bem*?"/"Can you give me the one from Mexico?").

Using this method, Diesendruck and Markson (2001) found that children made mutual exclusivity inferences for novel words regardless of whether Percey was present or absent during the labeling event. At first pass, these results suggest that children were not sensitive to speaker-specific knowledge when making mutual exclusivity inferences. Interestingly, however, Diesendruck and Markson

[1] It is possible that infants also deploy epistemic reasoning. Some have argued that infants as young as 6 months old can track what a speaker can see in order to determine their likely goals (Luo & Johnson, 2009; Sodian et al., 2007), and that 15- to 18-month-olds can represent that a speaker can have false beliefs about the location, identity, or properties of objects (Buttelmann et al., 2009; Scott & Baillargeon, 2009; Song & Baillargeon, 2008; Onishi & Baillargeon, 2005). However, there are also reports of failures (Sodian & Thoermer, 2008; Dunham et al., 2000; Poulin-Dubois et al., 2007) and non-replications of such findings (Powell et al., 2018; Dörrenberg et al., 2018; Kulke et al., 2018; see Lavelle, 2022 for a discussion).

found that children later judged that Percy knew the names of all labels, even when he was absent during the initial labeling session. To interpret these results, they argued that children might make a conventionality assumption (Csibra & Gergely, 2009; Henderson & Sabbagh, 2010), and assume that all competent speakers of a language share knowledge of words and their meanings (see also Diesendruck & Markson, 2011 for further discussion). Diesendruck and Markson took this as evidence that children do reason epistemically, because by adopting the conventionality assumption, they infer that Percy, in particular, knows the word 'dax' even if was absent when the child first learned it. This account aligns with a larger body of work investigating children's formation and understanding of conventionality, which finds that children as young as 9 months old make a conventionality assumption for labels (Henderson & Woodward, 2012; Graham et al., 2006; Buresh & Woodward, 2007), and even nonlinguistic knowledge such as objects' function (Casler & Kelemen, 2005) and rules of games (Rakoczy et al., 2008; Hardecker et al., 2016).

Might epistemic reasoning be involved in mutual exclusivity?

While the results in Diesendruck and Markson (2001) are compatible with the use of a conventionality assumption, it may not resolve the question of whether mutual exclusivity inferences involve epistemic reasoning. In particular, Srinivasan et al. (2019) note two factors that limit the interpretation of their findings. First, although it is possible that children made mutual exclusivity inferences even when Percy was absent due to a conventionality assumption, it is also possible that they instead reasoned egocentrically when reasoning about word meanings, and used their own knowledge of the object labels to guide their judgments. Second, Srinivasan et al. argued that Diesendruck and Markson may not have provided sufficiently strong evidence that Percy was ignorant of the novel labels used in the task. In particular, they argued that children might assume conventionality but fail to modulate this assumption in a speaker-specific way, in order to allow for the possibility that different speakers in the same linguistic community might have knowledge of different words. Compatible with this concern, previous studies find that when children are taught novel labels (such as 'dax'), they make a blanket assumption that other speakers of the language share these meanings (Csibra & Gergely, 2009; Henderson & Sabbagh, 2010). While conventionality is a useful heuristic to adopt in naturalistic conversations, it is also common for speakers of a language to not know all the words shared by the community (e.g., someone who does not play baseball might not know baseball-related terms). Therefore, proficient language users need to be able to modulate the conventionality assumption with reasoning about specific-speaker knowledge.

Given these concerns, Srinivasan et al. (2019) conducted a follow-up study using the paradigm of Diesendruck and Markson (2001), but manipulated how each novel label was introduced. In a Pedagogical condition, the experimenter taught children the name of one of two objects as in Diesendruck and Markson, while in a Coinage condition the researcher told the child that one object had no name, and asked them to help choose a label for it (e.g., "Look at this one! What should we call it? Should we call it a *dax* or a *zev*?"). As in Diesendruck and Markson, Percy was either present or absent during the labeling event, and emerged to make a request using a different label (e.g. 'Can you give me the *bem*?'). Therefore, Srinivasan et al. reasoned that when Percy was absent during coinage, participants should reason that he did not know its meaning, and therefore should not use it to inform mutual exclusivity inferences. However, against this, results from four experiments mostly supported a non-epistemic account. First, children (and adult controls) generally made mutual exclusivity inferences in both the Pedagogical and Coinage conditions, both when Percy was present and absent during labeling. Second, children's theory of mind ability did not reliably predict whether they made mutual exclusivity inferences. And finally, although both children and adults judged that Percy knew labels when they were taught or coined in his presence vs. but not when he was absent, this knowledge did not affect mutual exclusivity judgments, which were just as frequent when Percy was ignorant as when he was knowledgeable.

To interpret these results, Srinivasan et al. (2019) argued that mutual exclusivity inferences depend upon the assumption that speakers of a language will use conventional labels, but not upon the assumption that individual speakers actually know these words; a child who knows the word *spoon* might expect others to use this word to label spoons, and find it odd for them to use an alternative label, even if there's reason to doubt that they've heard it before. Consistent with this analysis, they found that both children and adults made fewer mutual exclusivity inferences when they were asked to make up a completely new name (Experiment 4), compared to when they were asked to pick between two labels suggested by the experimenter (Experiment 2). According to Srinivasan et al., this is expected because creating an entirely new word provides strong evidence that the new label is not conventional, whereas asking children to choose from between two labels leaves open the possibility that the choices are known by other speakers of the language – e.g., that they are playing a guessing game, to choose which is the 'correct' label used by other speakers. Thus, children may adopt a conventionality assumption even for the coined labels and expect others to observe these conventions even when they are not present during the coinage event.

In sum, previous accounts agree that children's mutual exclusivity inferences are supported by the conventionality assumption: upon hearing a novel label, they assume that other speakers in the same linguistic community share the same labels for objects. However, past accounts disagree about whether this assumption is mediated by epistemic reasoning. The hypothesis of Diesendruck and Markson (2001) suggests that children reason about interlocutors' knowledge while making conventionality assumptions. When children suspend this assumption, it is because they evaluate that a specific speaker does not share a particular piece of knowledge. Meanwhile, Srinivasan et al. (2019) propose that children make a conventionality assumption that extends to all speakers in the linguistic community without reasoning directly about individual epistemic states, and that they suspend this assumption when given evidence that a label is not conventional.

The present study

In the present study, we sought to adjudicate between these views. To do so, we revisited the concern raised by Srinivasan et al. (2019) that the Percy task in Diesendruck and Markson (2001) may not promote reasoning about epistemic states. In their study, Srinivasan et al. sought to address this issue by introducing a Coinage condition. However, in all of their experiments, as in Diesendruck and Markson, children were presented with only two objects, and then heard Percy use a word to request one of them. The problem with this is that Percy's use of a word implied that he knew its meaning, and therefore that it must refer to an object that has a conventionally known label – and not to the object that just received a newly coined label. Thus, the two-alternative forced choice paradigm may force children to rule out the object with the coined label, leading to a mutual exclusivity inference.

While one of the experiments conducted by Srinivasan et al. (2019) attempted to head off this concern by telling participants that both objects already had conventional labels, children may have been confused by this information. First, children were not explicitly told whether Percy knew these conventional labels, only that names existed. Second, perhaps amplifying this possible inference, in the Coinage condition children were explicitly told that the experimenter didn't know the labels. Thus, a truly epistemic reasoner might be confused about what conventional labels Percy should and should not know. Finally, even if children were capable of reasoning epistemically, they may have based their judgments on their own egocentric knowledge of labels anyway, simply because there was no other basis for making a choice between the two objects. Such an analysis would explain the finding that participants made mutual exclusivity inferences when Percy was absent during coinage, but also the fact that they made these inferences even when they acknowledged that Percy did not know the made-up word.

Here, we aimed to address these concerns and to assess the role of epistemic reasoning in preschoolers' mutual exclusivity inferences. First, in keeping with previous work, we first aimed to differentiate between the egocentric and the conventionality views by manipulating whether labels were taught or coined. Second, and new to this study, to avoid confusion about Percy's epistemic state we explicitly told children which labels Percy knew and did not know. Third, to address the concern that children might rely on heuristics created by a two-alternative forced choice, we added a third object. On each trial, Percy explicitly stated that he knew the names of only two objects, but not the third. Then, one of the objects Percy knew was labeled. Given this, when Percy subsequently requested an object children could not use the heuristic of picking the object that had gone unnamed (since there were two such objects), unless they integrated information about Percy's epistemic state (i.e., which objects he knew labels for). Fourth, when participants responded non-egocentrially – and computed mutual exclusivity inferences selectively in the Pedagogical condition but not the Coinage condition – we further probed the role of epistemic reasoning by asking participants if they thought Percy knew the label that was used in his absence. We reasoned that if mutual exclusivity inferences depend upon epistemic reasoning, then they should only compute inferences when they believe Percy knows the label, but not when he does not, compatible with an epistemic account.

Finally, we asked whether age and Theory of Mind ability were related to the use of epistemic information in the service of mutual exclusivity. Previous studies report that children between 3 and 6 years of age selectively learn novel words from speakers who use known words accurately, and that this ability is predicted by measures of Theory of Mind (e.g., Birch et al., 2008; Brosseau-Liard et al., 2015; Corriveau et al., 2011; Jaswal & Neely, 2006; Luchkina et al., 2020; see Tong et al., 2020; Sobel & Finiasz, 2020 for meta-analyses). These findings suggest that older children with more advanced Theory of Mind abilities may be more likely to base mutual exclusivity inferences on epistemic reasoning.

Method

A preregistration is available at https://osf.io/p9yg7. All materials, data and analysis code are available at https://osf.io/fwvbp.

Participants

Our final sample included 81 children (36F, 45 M; $M_{age} = 4.68$ [3.51; 5.95]; $SD_{age} = 0.57$) who were recruited from preschools in Comox, British Columbia, in Canada, as well as 80 adults[2] (49F, 31 M; $M_{age} = 44.28$ [20; 76]; $SD_{age} = 15.52$) recruited online from Prolific. Forty children participated in the Pedagogical condition (20F, 20 M; $M_{age} = 4.70$ [3.74; 5.96], $SD_{age} = 0.55$), and 41 children participated in the Coinage condition (16F, 25 M; $M_{age} = 4.67$ [3.51; 5.78], $SD_{age} = 0.59$). Forty adults participated in the Pedagogical condition (24F, 16 M; $M_{age} = 42.20$ [20;76], $SD_{age} = 14.91$) and 40 adults participated in the Coinage condition (25F, 15 M; $M_{age} = 46.35$ [21; 76], $SD_{age} = 16.02$). All participants spoke English as their primary language. All adult participants resided in the United States. This study was approved by the UCSD Human Research Protections Program (Protocol #171652).

Materials & Procedures

Following the procedure from Srinivasan et al. (2019), all children participated in a word learning task and a Theory of Mind battery (Wellman & Liu, 2004). Adult participants only completed the word learning task. Stimuli in the primary task were presented in a series of videos that depicted two actors sitting behind a table, and Percy, who was controlled by an experimenter who was hidden

[2] One additional adult was tested, but did not pass the initial memory check questions for five out of six trials despite being given three opportunities to respond per trial. Given that this only occurred in one trial across all other 80 participants, we took this as evidence that the participant did not demonstrate sufficient attention, and excluded them from the sample as pre-registered.

from view. There were 18 novel objects used in the word learning task (see Supplementary Materials S1). The novel labels were 12 monosyllabic, legal English words (*dax, lof, tig, gup, blick, mef, bem, bip, zot, kiv, wug, jop*).

Mutual exclusivity task

Participants were randomly assigned to one of two between-subject conditions: the Pedagogical condition and the Coinage condition. The main difference between the two conditions was how the novel label was introduced: in the Pedagogical condition, the novel label was 'taught' by one of the actors in the video, whereas in the Coinage condition, the two actors in the video invented the novel label.

Participants watched videos (see Supplementary Materials S2 for the full script), in which they were first introduced to two actors (a researcher and a confederate) sitting behind a table. The researcher introduced the puppet, Percy, describing that Percy was napping in his house and couldn't hear them: "I'll introduce you to my friend Percy. Percy lives in this house, but he can't hear us when he's inside the house." She asked the other actor to knock on the table to bring Percy out. Percy (a puppet controlled by an unseen third experimenter) then emerged from under the table, and introduced three of his toys (three novel objects). To emphasize that he could not hear from inside the house, Percy announced that he was tired and went back under the table for a nap. Then, the researcher asked him to come out but he did not. Participants were also asked why they thought Percy was not coming out. Regardless of their response, the video continued with the researcher then remembering that Percy could not hear from inside his house, and the confederate knocking on the table to bring him out.

Explicit knowledge phase. In each trial, Percy then emerged and approached each toy in sequence, saying whether he knew the name of each (making clear that he knew names for two of them, but not the third; see Fig. 1A). The location of the toy that Percy did not know the name of was counterbalanced between the leftmost and rightmost object. We avoided the middle object to keep the objects that he did know the name of together, to reduce memory load for children. Participants' memory of Percy's knowledge status was then probed by asking which objects he knew the labels and did not know the labels for. If they answered incorrectly, Percy described his knowledge of the toys again. If participants still made errors after three tries, the trial continued. Participants' performance on these memory check questions was generally high. Adults correctly reported Percy's knowledge for all objects on their first try on 96.25% of trials (462/480 trials), and by the third try on 99.79% of trials (479/480 trials). Children correctly reported Percy's knowledge for all objects on their first try on 83.47% of trials (404/484 trials), and by the third try on 96.49% of trials (467/484 trials; see Supplementary Materials, Table S1, for detailed results). Percy then announced that he was tired and wanted to sleep, and disappeared under the table again.

Labeling phase. One of the objects that Percy knew the name of was selected to be labeled – we refer to this as the Labeled & Known object. The objects that were not labeled will be referred to as the Unlabeled & Known and Unlabeled & Unknown objects, based on whether Percy reported knowing their names or not. In the Pedagogical condition, the researcher taught the participant the name of the Labeled & Known object (e.g., "Look at this one! I'll tell you what this one is called. It's a *dax*. This is a *dax*.") In the Coinage condition, the researcher emphasized that the toys had names but she did not know them, and that they would make up some names together. She then asked the confederate (the other actor in the video) to come up with a name for the Labeled & Known object: "It has a name but I don't know what it is. Hmm, what should we call it? Can you make up a name for it?" Once the name had been invented (e.g., "Let's call it a *dax*"), the researcher ratified the name (e.g., "Okay, let's call it a *dax*").

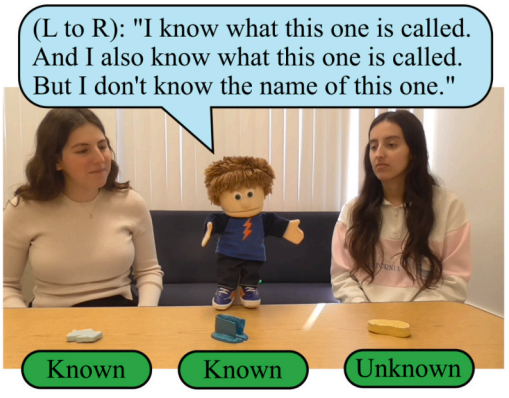
After each object was labeled, we asked participants whether Percy knew the label. In the Pedagogical condition, they were asked, e.g., "Does Percy know that this is a *dax*?" In the Coinage condition, they were asked, e.g., "Does Percy know that our friends call this a *dax*?" (for adult participants, the researcher and confederate in the video were introduced as Annie and Bonnie, so we asked, "Does Percy know that Annie and Bonnie call this a *dax*?"; Fig. 1B1-2). Participants' judgments of Percy's label knowledge were used to determine whether their referential choice was guided by epistemic reasoning about Percy's, in order to differentiate between the Epistemic account and the Non-Epistemic account. One child did not answer this question for three trials, so we excluded those trials for this analysis.

In both conditions, the actor also made neutral comments towards the Unlabeled & Known/Unlabeled & Unknown objects in order to equate attention towards each object. After each comment towards the unlabeled toys, participants were also probed for their memory of Percy's knowledge of that object (from his initial statements before leaving). Performance on these memory check questions was also high: adults answered both memory check questions correctly 91.88% of trials (441/480 trials), and children answered both correctly 74.38% of trials (360/484 trials; see Supplementary Materials, Table S1, for more detailed results). We did not exclude participants based on their responses to these memory check questions, but did repeat our main analyses with only participants with just the participants who succeeded on all memory check questions (including the initial memory check questions before the labeling).

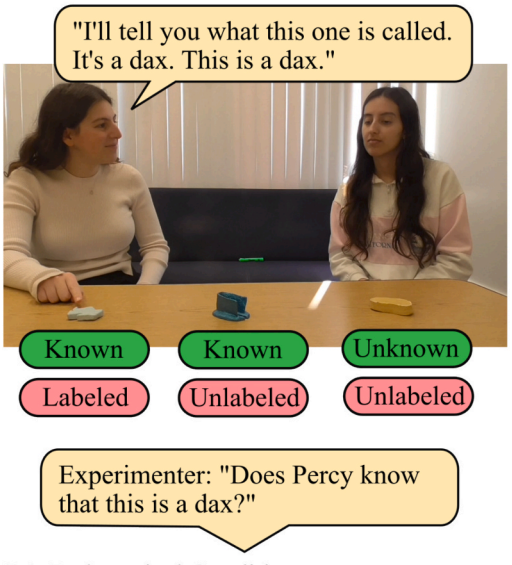
Request phase. After the labeling phase, Percy emerged from under the table and requested an object using a different novel label – e.g., "I want the *bem*. Can you give me the *bem*?" (Fig. 1C). Participants' choice of object was recorded and used for the main analyses. One child did not answer this question for one trial, so we excluded that trial for these analyses.

In total, participants saw six critical trials, in four pseudo-randomized orders. The order of how objects in each trial were physically presented (left to right) was counterbalanced ensuring that the location of the Unlabeled & Known object differed across trials. To aid memory, the objects that Percy knew the names of were always grouped together (so either the left and middle objects, or the middle and the right objects).

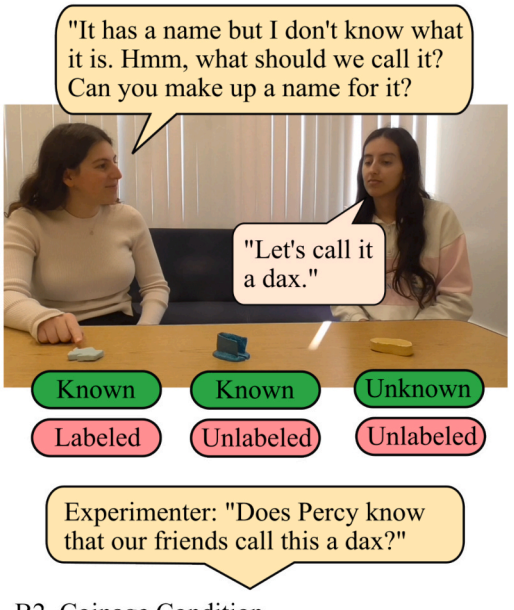
After the third trial, participants were reminded that Percy was unable to hear them from under the table. Finally, after all six critical trials, participants were asked, "Can Percy hear you from inside his house?" (following Srinivasan et al., 2019). The majority of



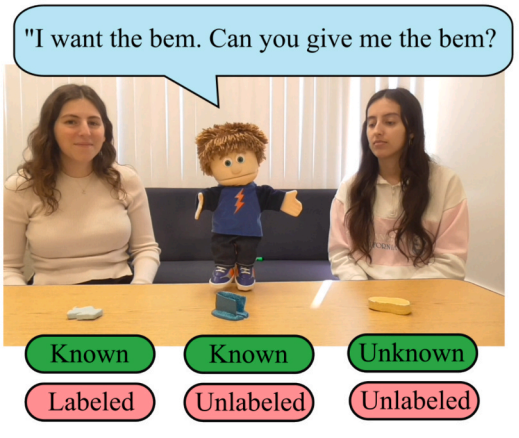
A. Percy states his knowledge status regarding the three objects, then leaves.



B1. Pedagogical Condition.



B2. Coinage Condition.



C. Percy returns and makes a request with a new label.

(caption on next page)

Fig. 1. *Schematic of study method. Note*. Images represent videos shown to participants. Each participant was randomly assigned to see only B1 or B2. Green labels indicate Percy's explicit knowledge/ignorance status, from A. Red labels indicate whether the object was named by the actors, from B1/2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

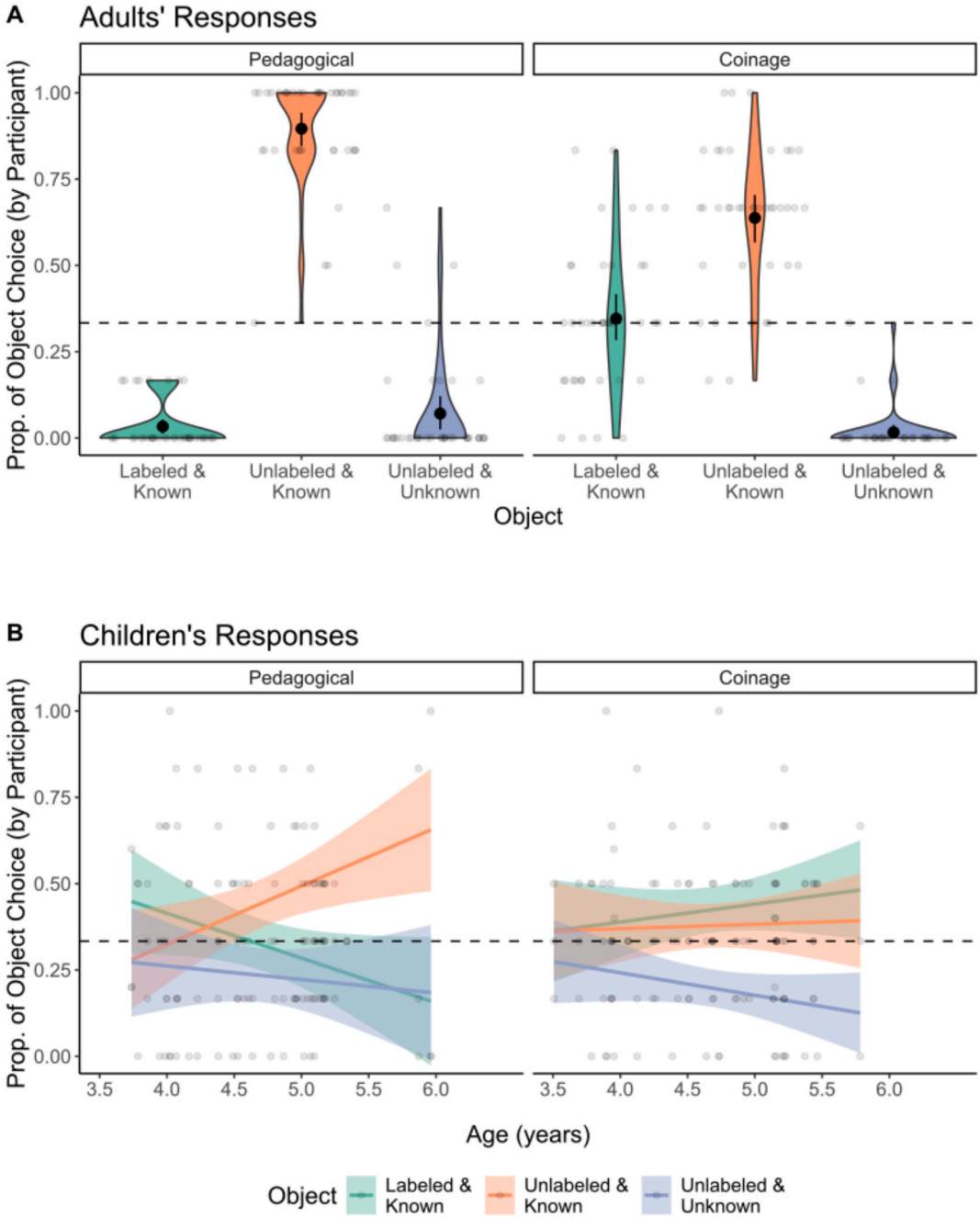


Fig. 2. *Adults and children participants' responses. Note*. Proportion of choice of each object for each condition, for A) adult participants (against object choice) and B) children participants (against age). Each dot represents a participant. Horizontal dashed line indicates selection at chance = $^1/_3$. Error bars and shaded areas represent 95% confidence intervals.

participants (78 adults, 97.50%; 67 children, 82.72%) responded 'no,' which was the expected response demonstrating that they understood the manipulation. This question aimed to probe participants' procedure comprehension, but we did not plan inferential analyses. Detailed results can be found in the Supplementary Materials, Table S1.[3]

Theory of Mind task

Following Srinivasan et al. (2019), each child received a battery of five items from Wellman and Liu's (2004) Theory of Mind questionnaire. In these trials, characters were represented with cardboard cutouts, and other items (e.g., containers) were represented physically. These questions tested their beliefs that: (1) others can hold different desires from theirs and will act on their own desires, (2) lack of visual access results in lack of knowledge about the contents of a container, (3) others can hold false beliefs about the contents of a container, (4) others can hold different beliefs from theirs and will act on their own beliefs, and (5) others will act on their own false beliefs (see Supplementary Materials S3 for a detailed description of each question). We used Wellman and Liu's criteria to compute a Theory of Mind score for each child, which ranged from 0 to 5, and children received a mean score of 3.05 (SD = 1.32; see Supplementary Materials, Table S2, for detailed results for each question).

Analytical approach

We conducted two sets of pre-registered analyses. The first asked whether participants responded egocentrically by testing whether responses differed between the Pedagogical and Coinage conditions. As described above, the Egocentric account predicts that participants should make mutual exclusivity inferences in both conditions, whereas the Conventionality accounts predicts that they only do so in the Pedagogical condition and not the Coinage condition. Although a non-egocentric pattern is compatible with the use of epistemic cues – since participants were required to attend to Percy's knowledge states to exclude the third, Unlabeled & Unknown object, we conducted a second, exploratory, analysis asking whether participants' choice of the Labeled & Known object was predicted by their judgment of Percy's knowledge of the labels.

Results

Mutual exclusivity inferences differed between labeling conditions

Our first set of analyses aimed to test whether children reasoned egocentrically by comparing mutual exclusivity inferences across the Pedagogical and Coinage conditions. If participants compute mutual exclusivity inferences then they should choose the Unlabeled & Known object more often than the Labeled & Known object, and if they reason egocentrically then they should do this even in the Coinage condition, when Percy cannot possibly know the label used by the experimenter. According to our pre-registered plan, we computed a Mutual Exclusivity Difference Score (ME Score) for each participant, which was defined as the difference in the proportion of selecting the Unlabeled & Known object vs. the Labeled & Known object across all trials. Thus, a higher ME Score corresponds to a higher rate of mutual exclusivity inferences. A participant who only chose the Labeled & Known object across all trials would have an ME Score of -1, while a participant who only chose the Unlabeled & Known object across all trials would have an ME Score of 1.[4] We then constructed a series of linear regression models to test whether labeling condition (dummy coded: Coinage = 0; Pedagogical = 1) predicted ME Score.

For adults, we found that Condition was a significant predictor of ME Score, driven by the fact that mutual exclusivity inferences were more frequent in the Pedagogical condition (M = 0.68, SD = 0.18) than in the Coinage condition (M = 0.29, SD = 0.42; $\beta_{Pedagogical} = 0.57$, 95% CI [0.43, 0.72], $t(78) = 7.87$, $p < .001$). In the child sample, Condition was not a significant predictor of ME Score (Pedagogical: M = 0.12, SD = 0.41; Coinage: M = −0.05; SD = 0.37; $\beta_{Pedagogical} = 0.16$, 95% CI [-0.01, 0.34], $t(79) = 1.90$, $p = .061$). In children, likelihood-ratio tests found that adding Age, or Age + ToM score as main effects in the regression did not improve the model (Model with Age: $\chi^2(1) = 0.35$, $p = .125$; Model with Age and ToM score: $\chi^2(2) = 0.38$, $p = .282$). Neither Age or ToM score was a significant predictor in both models ($ps > .05$). Thus, children, but not adults, appeared to reason egocentrically, and computed mutual exclusivity inferences even when the label used by the experimenter was coined and therefore not known to Percy.

An exploratory model found a significant interaction effect of Condition * Age ($\beta_{Pedagogical * Age} = 0.20$, 95% CI [0.03, 0.37], $t(76) =$

[3] Following Srinivasan et al. (2019) and Diesendruck and Markson (2001), we also probed participants' judgment of Percy's knowledge of the labels one more time at the end of the study (e.g., "Do you think Percy knows the names of the toys? / Do you think Percy knows what our friends called the toys?") However, given that Percy explicitly stated his knowledge of the toys at the start of each trial, and there were both toys he knew and did not know the name of, participants might have been confused about how they should answer this question. As such, their responses might not be informative, so we have reported the results in the Supplementary Materials, Table S1, instead of the main text.

[4] We did not consider Unlabeled & Unknown object choices in this score, since selection of this object is ambiguous between the child computing a mutual exclusivity inference (since this choice excludes the already labeled object) and a misunderstanding of the task (because Percy explicitly expressed ignorance of that object's name). Pre-registered *t*-tests and exploratory GLMMs showed that selection of this object is significantly less than expected by chance and did not differ across labeling conditions, age, or Theory of Mind abilities (see Supplementary Materials S4 for detailed statistics). Therefore, excluding choice of this object from calculating the ME Score should not significantly change the inferences from the following regression models.

2.36, $p = .021$), but no main effect of Condition, Age, or ToM ($p > .05$). A likelihood ratio test found that this model improved fit relative to the model with just a main effect of Condition ($\chi^2(3) = 1.18, p = .041$). A second exploratory model found that there was no significant Condition * ToM interaction ($\beta_{Pedagogical * ToM} = 0.11$, 95% CI [-0.07, 0.30], $t(76) = 1.23, p = .221$). Further, this model did not improve upon the Condition-only model ($\chi^2(3) = 0.61, p = .254$) in a likelihood ratio test. Finally, a model that contained a three-way interaction of Condition * Age * ToM Score found that only Condition and Condition * Age were significant predictors (Condition: $\beta_{Pedagogical} = 0.22$, 95% CI [0.04, 0.40], $t(73) = 2.39, p = .019$; Condition * Age: $\beta_{Pedagogical * Age} = 0.20$, 95% CI [0.01, 0.38], $t(73) = 2.07, p = .042$). Likelihood-ratio tests found that the three-way interaction model did not improve on either the Condition-only model ($\chi^2(6) = 1.62, p = .079$) or the Condition * Age model ($\chi^2(4) = 0.53, p = .449$). In sum, though overall children reasoned egocentrically, exploratory models found that older children were more likely to compute mutual exclusivity inferences in the Pedagogical condition than in the Coinage condition relative to younger children. However, variability in ToM abilities played no role in predicting differences.[56]

Evaluations of Percy's knowledge guided mutual exclusivity inferences

Our findings thus far indicate that adults and older children made more mutual exclusivity inferences when the label was taught (Pedagogical condition) compared to when it was coined (Coinage condition). These results suggested that they did not reason egocentrically while making mutual exclusivity inferences. Our second set of analyses aimed to further probe the role of epistemic reasoning in these inferences.

A critical point that differentiates between the epistemic and non-epistemic accounts is whether participants' response to Percy's request is predicted by their judgment of his knowledge of the label used by the experimenter. On the non-epistemic account, if participants assume that a taught label is conventional and a coined label is not, then we should find no relationship between their judgement of Percy's knowledge and their referential choice. Participants should make referential decisions based mainly on how the label was introduced: excluding the Labeled & Known object in the Pedagogical condition, and not excluding this object in the Coinage condition. However, on the epistemic account, we should find a relationship between participants' judgment of Percy's knowledge and their referential choice.

To test this question, we constructed a series of generalized linear mixed-models predicting whether participants chose the Labeled & Known object on each trial, with Label Knowledge Judgment and Condition as main effects, and random intercepts for each participant and each label used in Percy's request. Because we did not have priors as to which predictor would explain more variance in the data, we compared models that included each of the two predictors separately and a model that included both predictors as main effects. Model comparisons were done by comparing model fit statistics ΔAICc and ΔBIC. The models that best explained the data were those with only Label Knowledge Judgment as the main effect in both the adults and children sample (see Table 1). Both adults ("Percy knows": M = 0.03, SD = 0.18; "Percy doesn't know": M = 0.35, SD = 0.48) and children ("Percy knows": M = 0.33, SD = 0.47; "Percy doesn't know": M = 0.44, SD = 0.50) were significantly less likely to choose the Labeled & Known object when they judged that Percy knew the introduced label (adults: β_{yes} = -3.02, 95% CI [-3.86, −2.16], $p < .001$; children: β_{yes} = -0.51, 95% CI [-0.95, −0.07], $p = .024$). Adding Condition as an additional fixed effect did not improve the models' fit in a likelihood-ratio test ($p > .05$). Finally, when examining only the child data, we found that adding Age and Theory of Mind Score as main effects to the Label Knowledge Judgment-only model also did not improve the fit of the model in a likelihood-ratio test ($p > .05$). Thus, in summary, these results suggest that the effect of Condition (Pedagogical vs. Coinage) on mutual exclusivity inferences is explained in part by how participants reasoned about Percy's knowledge of the new label, consistent with an epistemic account.[7]

General discussion

We investigated how children and adults reason about speaker knowledge when making mutual exclusivity inferences. We found two main results that together provide evidence for epistemic reasoning among older children and adults, but egocentric reasoning in younger children. First, in a three-alternative forced-choice task where a speaker directly told participants his knowledge (and ignorance) of three objects' names, we found that adults made more mutual exclusivity inferences when a label was taught to them compared to when the label was invented. In contrast, we found that children were overall egocentric, but that this changed as a

[5] We repeated all analyses with both the planned and exploratory models on just trials where participants passed both memory checks (adults: 441/480 (91.88%) trials, N=78; children: 358/484 trials (73.97%) trials, N=72). We found similar results, except that in the model with the three-way interaction of Condition * Age * ToM in children, no predictors were statistically significant (*p*s > .05).

[6] Preregistered *t*-tests that analyzed selection of the Labeled & Known object (rather than the ME score) found the same pattern of results. Adults chose the Labeled & Known object significantly less often than chance ($^1/_3$) in the Pedagogical condition (M = 0.03, SD = 0.18, $t(39) = -28.10, p < .001$), and were at chance in the Coinage condition (M = 0.35, SD = 0.48, $t(39) = 0.37, p = .715$). Children selected the Labeled & Known object at chance for the Pedagogical condition (M = 0.32, SD = 0.47, $t(39) = -0.26, p = .793$), and significantly more often than expected by chance for the Coinage condition (M = 0.43, SD = 0.50, $t(40) = 2.75, p = .009$). Both condition effects were confirmed with exploratory GLMMs as reported in the main text. A follow-up model in children also found a significant Condition * Age interaction ($\beta_{Pedagogical * Age} = -0.51$, 95% CI [-0.97, −0.05], $z = -2.16, p = .031$), again suggesting that older children reason non-egocentrically.

[7] We note that Srinivasan et al. (2019) found that Label Knowledge Judgment was not a predictor for referential choice. However, all the models tested were complex models where Label Knowledge Judgment was one of the main effects, rather than a single main effect as the best model here. This could have resulted in the difference we found.

Table 1
Comparisons between models predicting choice of Labeled & Known object.

Model predicting choice of Labeled & Known object	Adults				Children			
	AICc	ΔAICc	BIC	ΔBIC	AICc	ΔAICc	BIC	ΔBIC
~ Condition	366.16	4.41	382.77	4.41	618.14	3.30	634.76	3.33
~ Label Knowledge Judgement	361.74	*	378.35	*	614.83	*	631.43	*
~ Condition + Label Knowledge Judgment	361.81	0.07	382.56	4.20	615.34	0.51	636.06	4.64

Note. * indicates the best model for each sample, which was used as the reference to calculate ΔAICc and ΔBIC.

function of age, and that older children performed non-egocentrically, like adults. This contrasts with findings of previous studies, which found that both adults and children failed to consider the speaker's knowledge of object labels when it was not explicitly mentioned, but was instead implicit in the difference between the Coinage and Pedagogical conditions (Diesendruck & Markson, 2001; Srinivasan et al., 2019). Second, we found that this effect of labeling context on mutual exclusivity inferences was explained by participants' evaluation of Percy's knowledge of the labels. Participants who thought that the speaker knew the familiar labels were more likely to make mutual exclusivity inferences. Together, these results suggest that older children and adults do not reason egocentrically when making mutual exclusivity inferences, and that they are able to condition inference on the speaker's individual epistemic states.

In contrast to previous studies (e.g., Diesendruck & Markson, 2001; Srinivasan et al., 2019), participants in our experiment were more likely to make mutual exclusivity inferences when the familiar label was taught, compared to when it was coined. The most likely explanation for this is that we provided explicit information regarding the speaker's knowledge of each object (i.e., Percy saying that he knows the names of some objects, but does not know the name of others), which was emphasized by memory check questions. In contrast, previous studies required participants to infer the speaker's knowledge from whether they were present or absent during the labeling phase. This may have mattered for two reasons. First, explicitly mentioning the speaker's knowledge likely made this information more salient. Second, participants may have taken the explicit mention of Percy's knowledge as a cue that this information was relevant to the task, and that they should reason specifically about what *Percy* knew, rather than relying on their own egocentric knowledge or assumptions of conventionality. Thus, while our study demonstrates that children and adults reason epistemically when this information is contextually salient, it also leaves open the possibility that they often overlook the knowledge states of speakers in naturalistic contexts, when such information is not explicitly mentioned or otherwise salient. As noted in the Introduction, previous studies find that when children learn new words, they assume that other speakers of the language will also know these meanings (Csibra & Gergely, 2009; Henderson & Sabbagh, 2010). Similarly, experimental work in adults suggests that language users assume that speakers are competent by default, and only override this assumption when provided with strong contextual evidence that they are ignorant (Bale et al., 2025). Future research could explore this question, and which contexts prompt adults and children to reason epistemically about novel labels.

Also new to this study was the finding that sensitivity to the contrast between taught and coined labels was predicted by awareness of speaker knowledge, supporting epistemic accounts. Specifically, children and adults who believed that the speaker knew the labels used by the experimenter were more likely to make mutual exclusivity inferences. For example, if children believed that Percy knew a coined label that was used to refer to an object in his absence, they were subsequently more likely to make a mutual exclusivity inference and assume that Percy's request must be directed at an alternative object, not the one labeled by the experimenter. This is important, because it supports the thesis that many children and adults reason about a speaker's individual knowledge of words to constrain mutual exclusivity inferences, and do not rely purely on a conventionality assumption – i.e., that words known by the community at large should be known by Percy. This result, combined with the finding that older children and adults differentiate between the Pedagogical and Coinage conditions, are consistent with the thesis that language users are able to restrict mutual exclusivity inferences by reasoning about the epistemic states of speakers.

Do young children consider speaker knowledge when making mutual exclusivity inferences?

While we found that adults and older children can draw on speaker knowledge to make mutual exclusivity inferences, we failed to find such evidence in younger children. Previous studies report mutual exclusivity inferences in children as young as 12 months of age (Markman et al., 2003; Golinkoff et al., 1992; Pomiechowska et al., 2021; Merriman & Stevenson, 1997; Merriman & Bowman, 1989; Halberda, 2003). When considered in light of the present findings, this raises the possibility that younger children do not consider speaker knowledge when reasoning about a novel label – even when a speaker's ignorance is explicitly stated. If so, this would suggest that epistemic reasoning may not be strictly constitutive of mutual exclusivity reasoning, but may play an auxiliary role, deployed especially under circumstances when the knowledge states of speakers are thought to differ from those of others. Such a view is compatible with past studies which find that other forms of pragmatic inference, like conversational implicature, are often made without reference to epistemic reasoning by young children (Barner et al., 2018) and individuals with ASDs (Hochstein et al., 2018), but are nevertheless sensitive to knowledge states in older children and adults when context emphasizes the contrast between ignorant vs. knowledgeable speakers (Bergen & Grodner, 2012; Goodman & Stuhlmüller, 2013; Kampa & Papafragou, 2020; Kampa et al., 2025). Future studies should explore this question, and whether the earliest forms of mutual exclusivity inference assume conventionality, or are instead based purely on children's own knowledge of object labels.

Our findings in younger children raise the possibility that younger children are sensitive to epistemic cues, but that the task used in

this literature is simply not sensitive to this knowledge. The Percy task, which we adopted here in order to understand the results of previous studies, is relatively complex and requires children to consider knowledge of labels from multiple perspectives (the speaker's explicit statement of knowledge, whether the label is taught or coined, and the speaker's absence from the labeling event). Compatible with this worry, children as old as 5 sometimes fail to compute mutual exclusivity inferences in this paradigm (Diesendruck & Markson, 2001; Srinivasan et al., 2019), which is much older than in many previous studies that do not also test epistemic reasoning (e.g., Markman & Wachtel, 1988; Markman et al., 2003). Also, previous research reports that children as young as 3 track speakers' ignorance and knowledge of labels, and can use this information to guide whether to learn novel words from those speakers (e.g., Birch et al., 2008; Corriveau et al., 2011; Jaswal & Neely, 2006; Luchkina et al., 2020; see Tong et al., 2020; Sobel & Finiasz, 2020 for meta-analyses). These factors suggest that complex tasks like this may underestimate the role of epistemic reasoning in mutual exclusivity, especially in younger children. Given this, future work should attempt to reconcile these studies with work on younger children by developing tasks that integrate mutual exclusivity inferences and epistemic reasoning in a way that is more sensitive to this ability in younger children. Testing this question in younger children is important, because most previous studies of mutual exclusivity do not consider the role of epistemic reasoning in mutual exclusivity inference. As noted in the Introduction, some studies ask whether children prefer to learn from reliable or knowledgeable speakers over less reliable sources – treating knowledge as a trait of the individual – but do not ask how a speaker's knowledge of specific words and objects modulates inference, in keeping with Gricean models of word learning.

When do children and adults reason about speaker-specific knowledge?

The present study suggests that older children and adults can integrate epistemic reasoning about specific speakers' knowledge of labels to guide mutual exclusivity inferences. However, even if learners are capable of conditioning mutual exclusivity inferences on epistemic information, it's likely that they often assume conventionality without additional epistemic reasoning. In general, when children are exposed to novel words, they are not directly told who knows them and who does not, though this does sometimes arise. For example, in child-directed speech, parents sometimes mark their ignorance or uncertainty of objects' labels either explicitly (e.g., "I don't know what this is") or implicitly (e.g., "You think it's a bike?"; Henderson & Sabbagh, 2010). Given that such explicit statements of knowledge and ignorance are relatively infrequent, if children did not assume conventionality they would need to wait for direct evidence that each individual speaker of a language knows each particular word. By making a blanket conventionality assumption, children can assume that words generally have the same meanings for all speakers, and that other interlocutors can understand them when they use words (see Sabbagh & Henderson, 2007, for a further discussion). Even for individuals with sophisticated Theory of Mind abilities, it remains more efficient to assume that everyone from a linguistic community has the same label knowledge, and to only reason about a specific speaker's epistemic states when given direct evidence to doubt that they are competent or knowledgeable.

How does epistemic reasoning about label knowledge develop?

Our findings leave open the question of how children develop the ability to epistemically reason about what words specific speakers know, and how they use this information to reason about the meaning of new labels. One possibility is that this skill draws on the same mechanisms used in traditional Theory of Mind tasks, such as understanding that different people can hold different beliefs. Supporting this idea, an exploratory analysis found that children's Theory of Mind scores predicted their ability to correctly evaluate whether Percy knew the label that had been introduced (see Supplementary Materials, S5). However, we found that Theory of Mind abilities do not predict children's mutual exclusivity inferences. Taken together, these findings suggest that in this age range, Theory of Mind may influence mutual exclusivity inferences only indirectly, and may be mediated by children's ability to judge a speaker's knowledge of labels. It is an open question for future research to explore the mechanisms through which children draw on Theory of Mind abilities to reason about others' linguistic knowledge and learn novel words in complex labeling contexts.

In summary, our study found that adults and older preschoolers can take into account the labels a speaker knows when making mutual exclusivity inferences. These findings suggest that reasoning about a speaker's epistemic state is a source of information that children can use to interpret the meanings of novel labels. This highlights the importance of epistemic reasoning in word learning, and opens the door for future research to explore how reasoning about speaker knowledge integrates with inferences about word meanings under different contexts and mediating factors.

CRediT authorship contribution statement

Khuyen N. Le: Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **David Barner:** Writing – review & editing, Supervision, Resources, Methodology, Conceptualization.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the participants, the children participants' parents, and preschool teachers and directors in Comox Valley, British Columbia, Canada. The authors also thank Jolene Nesnas and Gabriella Ruiz for their assistance with stimuli creation. The authors are grateful for the feedback from the members of the Language and Development Lab at the University of California, San Diego.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2026.106487.

Data availability

I have shared the links to my data/code in the manuscript.

References

Au, T. K., & Glusman, M. (1990). The principle of mutual exclusivity in word learning: To honor or Not to honor? *Child Development, 61*(5), 1474–1490. https://doi.org/10.1111/j.1467-8624.1990.tb02876.x

Bale, A., Noguchi, H., Rolland, M., & Barner, D. (2025). Competence by default: Do listeners assume that speakers are knowledgeable when computing conversational inferences? *Journal of Semantics, 42*(1–2), 39–55. https://doi.org/10.1093/jos/ffae016

Barner, D., Hochstein, L. K., Rubenson, M. P., & Bale, A. (2018). Chapter 14. Four-year-old children compute scalar implicatures in absence of epistemic reasoning. In K. Syrett, & S. Arunachalam (Eds.), *Trends in Language Acquisition Research* (Vol. 24, pp. 326–349). John Benjamins Publishing Company. https://doi.org/10.1075/tilar.24.14bar.

Bergen, L., & Grodner, D. J. (2012). Speaker knowledge influences the comprehension of pragmatic inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(5), 1450–1460. https://doi.org/10.1037/a0027850

Birch, S. A. J., Vauthier, S. A., & Bloom, P. (2008). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition, 107*(3), 1018–1034. https://doi.org/10.1016/j.cognition.2007.12.008

Bloom, P. (2000). *How children learn the meanings of words* (pp. xii, 300). The MIT Press.

Brosseau-Liard, P., Penney, D., & Poulin-Dubois, D. (2015). Theory of mind selectively predicts preschoolers' knowledge-based selective word learning. *British Journal of Developmental Psychology, 33*(4), 464–475. https://doi.org/10.1111/bjdp.12107

Buresh, J. S., & Woodward, A. L. (2007). Infants track action goals within and across agents. *Cognition, 104*(2), 287–314. https://doi.org/10.1016/j.cognition.2006.07.001

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition, 112*(2), 337–342. https://doi.org/10.1016/j.cognition.2009.05.006

Casler, K., & Kelemen, D. (2005). Young children's rapid learning about artifacts. *Developmental Science, 8*(6), 472–480. https://doi.org/10.1111/j.1467-7687.2005.00438.x

Clark, E. V. (1988). On the logic of contrast. *Journal of Child Language, 15*(2), 317–335. https://doi.org/10.1017/S0305000900012393

Clark, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language, 17*(2), 417–431. https://doi.org/10.1017/S0305000900013842

Corriveau, K. H., Pickard, K., & Harris, P. L. (2011). Preschoolers trust particular informants when learning new names and new morphological forms. *British Journal of Developmental Psychology, 29*(1), 46–63. https://doi.org/10.1348/2044-835X.002009

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Diesendruck, G. (2005). The principles of conventionality and contrast in word learning: An empirical examination. *Developmental Psychology, 41*(3), 451–463. https://doi.org/10.1037/0012-1649.41.3.451

Diesendruck, G., Carmel, N., & Markson, L. (2010). Children's sensitivity to the conventionality of sources. *Child Development, 81*(2), 652–668. https://doi.org/10.1111/j.1467-8624.2009.01421.x

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology, 37*(5), 630–641. https://doi.org/10.1037/0012-1649.37.5.630

Diesendruck, G., & Markson, L. (2011). Children's Assumption of the Conventionality of Culture. *Child Development Perspectives, 5*(3), 189–195. https://doi.org/10.1111/j.1750-8606.2010.00156.x

Dörrenberg, S., Rakoczy, H., & Liszkowski, U. (2018). How (not) to measure infant theory of mind: Testing the replicability and validity of four non-verbal measures. *Cognitive Development, 46*, 12–30. https://doi.org/10.1016/j.cogdev.2018.01.001

Dunham, P., Dunham, F., & O'Keefe, C. (2000). Two-year-olds' sensitivity to a parent's knowledge state: Mind reading or contextual cues? *British Journal of Developmental Psychology, 18*(4), 519–532. https://doi.org/10.1348/026151000165832

Gathercole, V. C. (1989). Contrast: A semantic constraint? *Journal of Child Language, 16*(3), 685–702. https://doi.org/10.1017/S0305000900010795

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology, 28*(1), 99–108. https://doi.org/10.1037/0012-1649.28.1.99

Goodman, N. D., & Stuhlmüller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science, 5*(1), 173–184. https://doi.org/10.1111/tops.12007

Graham, S. A., Stock, H., & Henderson, A. M. E. (2006). Nineteen-month-olds' understanding of the conventionality of object labels versus desires. *Infancy, 9*(3), 341–350. https://doi.org/10.1207/s15327078in0903_5

Grice, H. P. (1975). Logic and conversation. In D. Davidson (Ed.), *The logic of grammar* (pp. 64–75). Dickenson Pub. Co.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition, 87*(1), B23–B34. https://doi.org/10.1016/S0010-0277(02)00186-5

Hardecker, S., Schmidt, M. F. H., Roden, M., & Tomasello, M. (2016). Young children's behavioral and emotional responses to different social norm violations. *Journal of Experimental Child Psychology, 150*, 364–379. https://doi.org/10.1016/j.jecp.2016.06.012

Henderson, A. M. E., & Sabbagh, M. A. (2010). Parents' use of conventional and unconventional labels in conversations with their preschoolers. *Journal of Child Language, 37*(4), 793–816. https://doi.org/10.1017/S0305000909990122
Henderson, A. M. E., & Woodward, A. L. (2012). Nine-month-old infants generalize object labels, but not object preferences across individuals. *Developmental Science, 15*(5), 641–652. https://doi.org/10.1111/j.1467-7687.2012.01157.x
Hochstein, L., Bale, A., & Barner, D. (2018). Scalar Implicature in Absence of Epistemic reasoning? the Case of Autism Spectrum Disorder. *Language Learning and Development, 14*(3), 224–240. https://doi.org/10.1080/15475441.2017.1343670
Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language Experience Shapes the Development of the Mutual Exclusivity Bias. *Infancy: The Official Journal of the International Society on Infant Studies, 15*(2), 125–150. https://doi.org/10.1111/j.1532-7078.2009.00009.x
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, 17*(9), 757–758. https://doi.org/10.1111/j.1467-9280.2006.01778.x
Kampa, A. F., Papafragou, A., & Jasińska, K. K. (2025). Scalar inference is supported by Theory of mind networks in adults and children. *Language, Cognition and Neuroscience, 1–21*. https://doi.org/10.1080/23273798.2025.2489602
Kampa, A., & Papafragou, A. (2020). Four-year-olds incorporate speaker knowledge into pragmatic inferences. *Developmental Science, 23*(3), Article e12920. https://doi.org/10.1111/desc.12920
Koenig, M. A., Clément, F., & Harris, P. L. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science, 15*(10), 694–698. https://doi.org/10.1111/j.0956-7976.2004.00742.x
Koenig, M. A., & Harris, P. L. (2007). The basis of epistemic trust: Reliable testimony or reliable sources? *Episteme, 4*(3), 264–284. https://doi.org/10.3366/E1742360007000081
Kulke, L., Reiß, M., Krist, H., & Rakoczy, H. (2018). How robust are anticipatory looking measures of theory of mind? Replication attempts across the life span. *Cognitive Development, 46*, 97–111. https://doi.org/10.1016/j.cogdev.2017.09.001
Lavelle, J. S. (2022). When a Crisis Becomes an Opportunity: The Role of Replications in making Better Theories. *The British Journal for the Philosophy of Science, 73*(4), 965–986. https://doi.org/10.1086/714812
Liittschwager, J. C., & Markman, E. M. (1994). Sixteen- and 24-month-olds' use of mutual exclusivity as a default assumption in second-label learning. *Developmental Psychology, 30*(6), 955–968. https://doi.org/10.1037/0012-1649.30.6.955
Luchkina, E., Corriveau, K. H., & Sobel, D. M. (2020). I don't believe what you said before: Preschoolers retrospectively discount information from inaccurate speakers. *Journal of Experimental Child Psychology, 189*, Article 104701. https://doi.org/10.1016/j.jecp.2019.104701
Luo, Y., & Johnson, S. C. (2009). Recognizing the role of perception in action at 6 months. *Developmental Science, 12*(1), 142–149. https://doi.org/10.1111/j.1467-7687.2008.00741.x
Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science, 14*(1), 57–77. https://doi.org/10.1207/s15516709cog1401_4
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
Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology, 47*(3), 241–275. https://doi.org/10.1016/S0010-0285(03)00034-3
Mather, E., & Plunkett, K. (2011). Mutual exclusivity and phonological novelty constrain word learning at 16 months. *Journal of Child Language, 38*(5), 933–950. https://doi.org/10.1017/S0305000910000401
Merriman, W. E., & Bowman, L. L. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development, 54*(3–4), 130. https://doi.org/10.2307/1166130
Merriman, W. E., & Stevenson, C. M. (1997). Restricting a familiar name in response to learning a new one: Evidence for the mutual exclusivity bias in young two-year-olds. *Child Development, 68*(2), 211–228. https://doi.org/10.2307/1131846
Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*(5719), 255–258. https://doi.org/10.1126/science.1107621
Pomiechowska, B., Bródy, G., Csibra, G., & Gliga, T. (2021). Twelve-month-olds disambiguate new words using mutual-exclusivity inferences. *Cognition, 213*, Article 104691. https://doi.org/10.1016/j.cognition.2021.104691
Poulin-Dubois, D., Sodian, B., Metz, U., Tilden, J., & Schoeppner, B. (2007). Out of Sight is not out of mind: Developmental changes in infants' understanding of visual perception during the second year. *Journal of Cognition and Development, 8*(4), 401–425. https://doi.org/10.1080/15248370701612951
Powell, L. J., Hobbs, K., Bardis, A., Carey, S., & Saxe, R. (2018). Replications of implicit theory of mind tasks with varying representational demands. *Cognitive Development, 46*, 40–50. https://doi.org/10.1016/j.cogdev.2017.10.004
Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology, 44*(3), 875–881. https://doi.org/10.1037/0012-1649.44.3.875
Sabbagh, M. A., & Henderson, A. M. E. (2007). How an appreciation of conventionality shapes early word learning. *New Directions for Child and Adolescent Development, 2007*(115), 25–37. https://doi.org/10.1002/cd.180
Sabbagh, M. A., & Shafman, D. (2009). How children block learning from ignorant speakers. *Cognition, 112*(3), 415–422. https://doi.org/10.1016/j.cognition.2009.06.005
Scofield, J., & Behrend, D. A. (2008). Learning words from reliable and unreliable speakers. *Cognitive Development, 23*(2), 278–290. https://doi.org/10.1016/j.cogdev.2008.01.003
Scott, R. M., & Baillargeon, R. (2009). Which Penguin is this? Attributing False Beliefs about Object Identity at 18 months. *Child Development, 80*(4), 1172–1196. https://doi.org/10.1111/j.1467-8624.2009.01324.x
Sobel, D. M., & Finiasz, Z. (2020). How Children Learn from others: An Analysis of Selective Word Learning. *Child Development, 91*(6), e1134–e1161. https://doi.org/10.1111/cdev.13415
Sobel, D. M., & Kushnir, T. (2013). Knowledge matters: How children evaluate the reliability of testimony as a process of rational inference. *Psychological Review, 120*(4), 779–797. https://doi.org/10.1037/a0034191
Sodian, B., & Thoermer, C. (2008). Precursors to a Theory of mind in infancy: Perspectives for Research on Autism. *Quarterly Journal of Experimental Psychology, 61*(1), 27–39. https://doi.org/10.1080/17470210701508681
Sodian, B., Thoermer, C., & Metz, U. (2007). Now I see it but you don't: 14-month-olds can represent another person's visual perspective. *Developmental Science, 10*(2), 199–204. https://doi.org/10.1111/j.1467-7687.2007.00580.x
Song, H., & Baillargeon, R. (2008). Infants' reasoning about others' false perceptions. *Developmental Psychology, 44*(6), 1789–1795. https://doi.org/10.1037/a0013774
Srinivasan, M., Foushee, R., Bartnof, A., & Barner, D. (2019). Linguistic conventionality and the role of epistemic reasoning in children's mutual exclusivity inferences. *Cognition, 189*, 193–208. https://doi.org/10.1016/j.cognition.2019.04.001
Tomasello, M. (2000). The social-pragmatic theory of word learning. *Pragmatics, 10*(4), 401–413. https://doi.org/10.1075/prag.10.4.01tom
Tong, Y., Wang, F., & Danovitch, J. (2020). The role of epistemic and social characteristics in children's selective trust: Three meta-analyses. *Developmental Science, 23*(2), Article e12895. https://doi.org/10.1111/desc.12895
Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-mind Tasks. *Child Development, 75*(2), 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x