

Seeing vs. Seeing That: Children’s Understanding of Direct Perception and Inference Reports

E. Emory Davis & Barbara Landau*

Abstract. Young children can reason about direct and indirect visual information, but fully mapping this understanding to linguistic forms encoding the two knowledge sources appears to come later in development. In English, perception verbs with small clause complements (“I saw something happen”) report direct perception of an event, while perception verbs with sentential complements (“I saw that something happened”) can report inferences about an event. In two experiments, we ask when 4-9-year-old English-speaking children have linked the conceptual distinction between direct perception and inference to different complements expressing this distinction. We find that, unlike older children or adults, 4-6-year-olds do not recognize that *see* with a sentential complement can report visually-based inference, even when syntactic and contextual cues make inference interpretations highly salient. These results suggest a prolonged developmental trajectory for learning how the syntax of perception verbs like *see* maps to their semantics.

Keywords. language acquisition; perception verbs; inference; semantics; syntax; complements

1. Introduction The use of perception verbs like *see* or *hear* with different complement structures often corresponds to reporting distinct kinds of perceptual experiences. In English, perception verbs with small clause complements can only report direct perception of an event as it occurred, while perception verbs with embedded clause complements can be used to report an inference about an event without having directly perceived it, as demonstrated by (1) and (2) below:

- (1) John saw the book fall off the shelf.
- (2) Mary saw that the book had fallen off the shelf.

(1) is true only if John witnessed the event of the book falling off the shelf. While (2) can be true for Mary in that same situation, it can also be true if Mary simply walked into the room after the book fell and noticed it on the floor next to the shelf.

This distinction relates to source monitoring, the ability to reflect on and distinguish between various sources of information and knowledge. Differences in perceptual representations and knowledge arise not just from having different experiences with objects, but also from having different access to events as they unfold. An individual like John in (1), who directly witnesses an event as it occurs, has a different perceptual experience from an individual like Mary in (2), who may have only seen the outcome of that event; even if John and Mary end up with similar representations of an event of a book falling off a shelf, John’s representation is based on directly witnessing the event, while Mary’s is based on an inference.

There is evidence that even young children can distinguish between direct and indirect sources of perceptual information. Ünal and Papafragou (2019) used two picture-matching tasks to assess

* We would like to thank research assistants Arunima Vijay, Sydney Sappenfield, and Clara Darcy for their work in making these experiments possible, as well as the members of the Language and Cognition Lab at Johns Hopkins for their insightful and valuable feedback at all steps of the research process. Authors: E. Emory Davis, Johns Hopkins University (emorydavis@jhu.edu) & Barbara Landau, Johns Hopkins University (landau@jhu.edu).

Turkish- and English-speaking children's ability to reason about perception and inference as sources of knowledge about events for both themselves and others. They found that 4-6-year-olds could use both direct visual evidence (a picture of a woman drinking for *drink*) and indirect visual evidence (a picture of footprints in the snow for *walk*) to reason about events and match photographs to verbs describing the events. Specifically, in the latter example, children needed to infer the occurrence of a past event in order to connect a verb like *walk* to a picture of footprints; they were also able to ascribe that same reasoning to someone else, though children performed less well when attributing either direct or inferential knowledge to others. Ünal and Papafragou's results stand in contrast to previous work that showed children under age 6 have difficulty with recognizing inference as a source of knowledge (Pillow 1999, Sodian & Wimmer 1987, *inter alia*). Ünal and Papafragou suggest that children's success in their tasks could be due to the fact that children were not required to produce or comprehend explicit verbal reports about visual access or mental states.

While these results indicate that 4-6-year-olds may be able to reason about perception and inference as sources of knowledge, there is also evidence that at this age children have difficulty mapping this understanding to linguistic forms that encode the difference between perception and inference, especially in comprehension. The term "evidential strategies" (Aikhenvald 2014) refers to the various systems languages use to mark knowledge sources, such as perception and inference. Despite the differences between languages in the types of evidential strategies they employ, the acquisition process for them appears to be remarkably similar cross-linguistically. Research has consistently shown that until around the age of six or seven, children's command of these strategies is not fully adult-like (e.g. in Tibetan; de Villiers et al. 2009), and children's comprehension lags behind their production (Papafragou et al. 2007, Ünal & Papafragou 2016, Ünal & Papafragou 2018, Winans et al. 2015). This asymmetry is notable not only for its robustness across languages and evidential typologies, but also because it is the reverse of many other production-comprehension asymmetries in language – children often understand linguistic forms before they produce them (Goldin-Meadow, Seligman & Gelman 1976, Gertner, Fisher & Eisengart 2006).

English marks the difference between direct perception and inference in perceptual reports using several different (optional) syntactic strategies. One way is through verbal complements of perception verbs like *see*. As discussed above, perception verbs with small clause complements (e.g. *John saw Mary leave*) report direct perception of an event, while perception verbs with embedded clause complements (e.g. *John saw that Mary left*) can be used to report either direct perception or inference about an event on the basis of visual (or sometimes other) evidence. Children face two potential challenges in learning that *see* can sometimes mean "infer." First, *see* + sentential complement maps to two different readings – direct perception or inference – so children must figure out which one of these is appropriate in the context; they must also learn that other frames, like small clause complements, do not allow an inference interpretation. Another aspect of this first challenge is that the direct perception reading may well be the primary or more accessible reading for the verb *see*. Second, learning the inferential meaning of *see* requires children to abstract away from the visual perception component of the verb's semantics to some extent. That is, *see* can report a belief one has about one thing (such as a book falling) as a result of perceiving of something else with entirely different visual properties (such as a book on the floor). Previous research has shown that young children (around age 4) do in fact have difficulty learning some aspects of perception verb semantics, and tend to assign narrower meanings to perception verbs than older children or adults, treating verbs like *see* as only reporting an event

involving direct perception with one’s eyes (Landau & Gleitman 1985, Elli, Bedny & Landau 2021).

In two experiments, we sought to determine whether and when young English-speaking children, who already produce perception verbs like *see* in both small clause and sentential complement frames (Davis & Landau 2020), have linked the conceptual distinction between direct perception and inference to the different complements expressing this distinction.

2. Experiment 1. In Experiment 1, we sought to determine whether 4-9-year-old children have mapped *see* in different syntactic frames to distinct kinds of perceptual experiences, and whether they recognize that perception verb utterances containing sentential complements can be true in different contexts than those containing small clause complements.

2.1. METHODS. We presented 36 children (4;0-9;01, $M = 6;5$) and six adults with eight different stories in which a first character directly perceives an event, while a second encounters visual evidence that could lead to an inference about the event. For example, in one of the stories, two children, Lily and Noah, leave a plate of cookies in the kitchen while they go out to play. Lily comes back inside and catches her dog Fido eating the cookies; Fido runs out of the room and Lily chases after him. Noah then comes into the kitchen and finds an empty plate on the floor surrounded by cookie crumbs and paw prints. All stories were of similar length and complexity. Narration of the stories and the sentences presented in the test phase (see below) were pre-recorded. The narrator did not use different voices for any of the characters. The narration was accompanied by visual depictions of the stories (Figure 1).

		
<p>This is Lily and Noah. Lily and Noah are friends. One day Lily and Noah were at Lily’s house and decided to bake some cookies. While they were baking, Lily’s dog Fido was taking a nap on his bed in the kitchen.</p>	<p>When the cookies came out of the oven, they were too hot to eat right away.</p>	<p>Lily and Noah put the cookies on a plate and left them on the counter in the kitchen.</p>
		
<p>Then Lily and Noah went outside to ride their bikes while the cookies cooled down. After riding bikes for a while, Lily went inside to get something to drink.</p>	<p>When Lily came into the kitchen, her dog Fido was eating the cookies that Lily and Noah made! Lily yelled, “Fido, stop!” and he ran away, and Lily ran after him.</p>	<p>Then Noah went inside to get a drink, too. Noah went into the kitchen and found the empty plate. All that was left were cookie crumbs and paw prints.</p>

Figure 1. Images and transcript of narration from one of the stories in Experiment 1.

After each story, participants heard two sentences: a Direct Perception sentence reporting perception of the event as it occurred (e.g. “I saw Fido eating the cookies”) and an Inference sentence reporting an inference based on visual evidence (e.g. “I saw that Fido had eaten the cookies”). They were then asked to identify which of the two characters (e.g. Lily or Noah) said *either* the Direct Perception (4 trials) or Inference sentence (4 trials). Participants also heard two pairs of control sentences for each trial, checking for their ability to identify the characters (e.g., “I have red pants”) and remember details of the story (e.g., “I came into the kitchen and found an empty plate”). There were two presentations of the task; each had the stories in a different randomized order and both counterbalanced the order of the sentences for each pair of targets and controls.

Both target sentences were presented in every trial as we thought that children might need the contrast of the two frames within each trial to be successful in linking the Inference sentence to the inferring character. That is, while *I saw that Fido had eaten the cookies* can be truthfully said by either the direct perception character (Lily) or the inferring character (Noah), the inference interpretation is strengthened by contrasting the embedded clause frame with the small clause frame, since *I saw Fido eat the cookies* can only be truthfully said by the direct perception character. Responses were marked correct if participants attributed the Direct Perception sentence to the direct perception character (on the 4 trials where this was queried) and the Inference sentence to the inferring character (on the other 4 trials), which was the expected pattern for adults.

2.2. RESULTS. Participant performance was measured as the proportion of correct responses for each sentence type (Direct Perception or Inference) across all stories. Adults performed as expected and attributed the Direct Perception sentences to the direct perception characters and the Inferences sentence to the inferring characters. Adult performance was at ceiling (mean correct above 95%) for both sentence types, so detailed analysis of their responses was not conducted. Both adults and children performed at ceiling for the control sentences.

Children’s performance was compared to chance (0.5 for each sentence type) using one-tailed single sample *t*-tests. Responses were also analyzed with logistic mixed effects models using the *lme4* package in R (Bates et al. 2014). These models included sentence type (Direct vs. Inference), age, gender, task presentation, and trial number as fixed effects, and participant and story as random effects. If a model did not converge with all of these fixed effects, gender was removed, then task presentation and trial number. Except for sentence type and age, none of these effects were found to be significant in any of the models that included them.

Children consistently attributed Direct Perception sentences to the direct perception characters ($M = 0.91$, $t(35) = 14.407$, $p < 0.05$), with no effect of age ($\beta = 0.02$, $SE = 0.02$, $p > 0.05$). In contrast, there was an age effect for the Inference sentences ($\beta = 0.11$, $SE = 0.03$, $p < 0.05$): children under 7 years of age ($n = 23$, $M = 5;5$) were at chance in attributing Inference sentences to the inferring character ($M = 0.55$; $t(22) = 0.8941$, $p > 0.05$), whereas children 7 and older ($n = 13$, $M = 8;4$) performed significantly above chance on the Inference sentences ($M = 0.96$, $t(12) = 17.725$, $p < 0.05$) and were no different from adults ($M = 0.96$).

We also examined joint performance on the two sentence types to determine whether individual children demonstrated a tendency to interpret both target sentences as having a direct perception meaning. Each child’s responses were categorized as fitting one of four patterns: above chance for both sentence types ($n = 22$), which was the adult pattern; above chance for Direct Perception and below chance for Inference ($n = 10$), which would indicate a direct perception interpretation of both targets; below chance for Direct Perception and above chance for Inference ($n = 2$); and below chance for both ($n = 2$). A multinomial logistic regression analysis showed that

age was a significant predictor of the two dominant response patterns ($\beta = -0.97$, $SE = 0.4$, $p < 0.05$; Figure 2), confirming that the older children understood the distinction between *see* and *see that*, while the younger children did not.

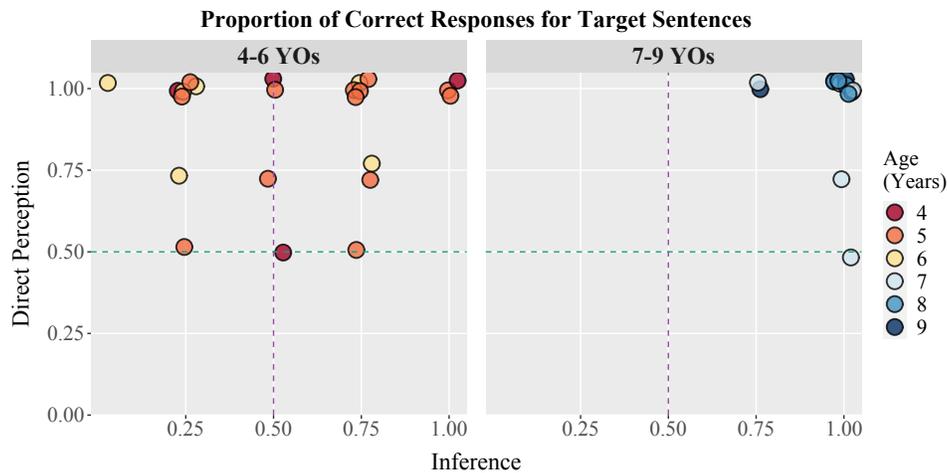


Figure 2. Proportion of correct responses for Direct Perception vs. Inference sentences in Experiment 1. Each point represents one participant; dotted lines show chance levels. Points in upper right corners indicate distinct interpretations for the Direct Perception and Inference sentences; points in upper left corners indicate a direct perception interpretation for both sentence types.

2.3. DISCUSSION. While children over 7 years old performed like adults on this task, and consistently attributed the Inference sentences to the inferring characters, children under 7 attributed the Inference sentences to the direct perception characters about half the time on average. The high rate of correct responses for the control questions indicates that participants of all ages understood the task and could follow the stories. Participants also rarely made errors on the Direct Perception statements, even if they responded incorrectly for the Inference statements. Any difficulty that the younger participants had with the Inference sentences, then, could not be due to issues with remembering the characters or the events in the stories. We believe there are two possible explanations for the younger children’s failure to consistently attribute the Inference sentences to the inferring characters: children’s linguistic knowledge or pragmatic factors.

First, younger children may have difficulty with the Inference sentences because they have incomplete knowledge of the semantics and syntax of perception verbs. Children who performed poorly on the Inference sentences may have believed that *see* in any syntactic frame can only refer to direct perception. This account fits with previous research which has shown that younger children assign narrow meanings to perception verbs, treating *see* as only referring to perceiving with one’s eyes. One 6-year-old participant who chose the direct perception character for every Inference sentence insisted on each trial that both target sentences were said by the direct perception character, suggesting that at least some children may have had this more limited semantic representation for *see*.

A second possibility is that younger children’s responses were driven largely by pragmatic considerations, in particular the salience of the direct perception character for them, rather than their knowledge of the syntactic frames. Given that the target sentences were statements about who saw what, children may have assumed that the experimenter wanted to know about the character who actually witnessed the event. Alternatively, children may have ranked direct perception higher

than inference in the set of possible interpretations for *see*. Participants may have believed that the direct perception character was more likely to be the answer since that character had more direct perceptual experience with the event than the inferring character. This could explain why younger children's responses skewed towards attributing both types of target sentences to the direct perception characters and almost never the inverse. Additionally, the fact that the *see* + embedded clause sentence is acceptable for the direct perception character to say may have further strengthened children's assumption that the character with direct visual experience was the better of the two choices. This contrasts with adults and older children, who did not make the direct perception interpretation of *see* with a sentential complement when it was contrasted with the small clause frame. These considerations may have given greater weight to the direct perception character in the younger children's decision process, overriding consideration of the syntactic distinction and the implicature pragmatics that accompany the contrast of the two sentence types.

Our results suggest that it is not until around age seven that English-speaking children consistently make adult-like distinctions between the syntactic frames that *see* occurs with and the corresponding semantics – that is, knowing that “I saw something happen” is different from “I saw that something happened” and that such statements are appropriate in different contexts. However, since direct perception and inference were both potentially acceptable readings of *see* + sentential complement in this task, we cannot determine whether younger children's difficulty is due to their understanding of the pragmatics of two readings, or because they do not have sufficient semantic and syntactic knowledge of *see*. The next experiment attempts to distinguish between these two possibilities.

3. Experiment 2. In Experiment 2, we tested whether younger children would accept *see that* for reporting inference in a truth-value judgment task designed to reduce some of the pragmatic complexities of Experiment 1. The TVJT provides participants with the opportunity to make independent judgments about each of the target sentences; children only need to determine whether *see that* is acceptable in an inference scenario, rather than identify the best or most likely interpretation as they may have done in the previous forced-choice task. This new task also depicts only one individual whose perceptual access to the event varies across trials, rather than two individuals with different perceptual experiences in each trial, eliminating the possibility that participants would implicitly compare different individuals' perceptual experiences and make linguistic judgments on the basis of who saw the event “better” (i.e. more directly). If children are aware that *see* + sentential complement licenses an inference reading, they should judge Inference sentences as “right” in inference scenarios while judging Direct Perception sentences as “wrong”; if they think *see* (in either frame) can only report direct perception, they should reject both target sentences when the speaker has only seen evidence.

3.1. METHODS. Participants were 14 adults and 23 children (4;05-6;10; $M = 5;06$). We included only children under 7 in this experiment since children 7 and older in Experiment 1 performed like adults.

The truth-value judgment task had two within-subjects factors: visual access to an event and sentence type. Visual access had three conditions (See Event, See Evidence, and Doesn't See) and each was tested with queries about two target sentence types (Direct Perception and Inference). Six different events were presented per condition for a total of 18 trials, and in every trial both Target sentence types were tested, plus one control sentence. The 18 target trials were presented to participants in one of three pre-determined randomized orders, with the orders themselves assigned randomly to participants. Only one version of each event was used per presentation order.

The order of the three test sentences (two Targets plus one Control) in each trial was randomized within each presentation of the task.

Participants were shown videos depicting an observer, Mary, watching a simple event in which an actor causes a visible change of state to an object (e.g. the actor peels a banana). Mary is prompted by the sound of a bell to put on a blindfold at different points during the videos, varying by the three visual access conditions (Figure 3). In the See Event condition (6 trials), Mary sees the entire event. In the See Evidence condition (6 trials), Mary sees the object beforehand and evidence of the event afterwards (the peeled banana), but not the peeling event itself. In the Doesn't See condition (6 trials), Mary sees the object before the event, but does not see the event or any evidence of it. Participants themselves always saw the full event, including how much of the event Mary watched.

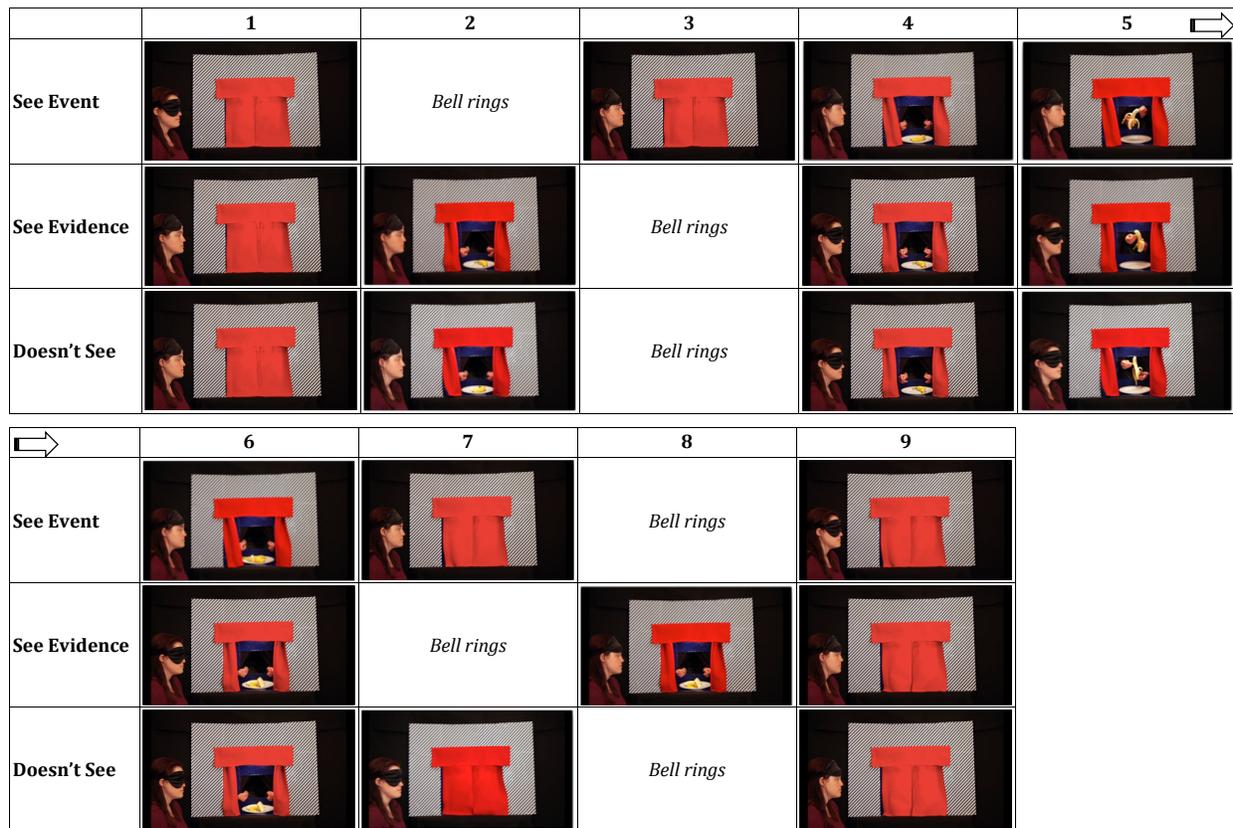


Figure 3. Timeline of videos in the three visual access conditions, showing when the observer, Mary, put on and removed her blindfold relative to the event.

After each event video, participants watched as Mary made three statements in separate videos: a Direct Perception statement (e.g. “I saw someone peel the banana”), an Inference statement (e.g. “I saw that someone peeled the banana”), and a control statement that could be true or false (e.g. “There was a banana” or “There was an orange”). Participants were asked if Mary was “right” or “wrong” after each sentence. The complements in the Direct Perception and Inference sentences always matched the event, so that participants would judge the statements based on Mary’s perception of the event rather than the felicity of the complement.

3.2. RESULTS. The expected adult-like responses to the target sentences for each visual access condition were based on adult performance in Experiment 1, as well as another experiment not reported here, which showed that under some circumstances, adults accept both *see* and *see that* for direct perception events.¹ Participants with adult-like knowledge of the semantics of *see* were expected to judge Direct Perception sentences (“I saw...”) as “right” only when Mary saw the event directly (See Event trials); to judge Inference sentences (“I saw that...”) as “right” when Mary either saw the event (See Event trials) or saw evidence of it (See Evidence trials); and to judge both target sentences as wrong when Mary did not see any aspect of the event (Doesn’t See trials). We also expected participants to judge true control sentences as “right” and false ones as “wrong.” Participants’ responses were coded as correct if they fit this pattern, and as incorrect if they did not. We expected that children who did not understand that *see that* can report inference would differ from this pattern in just one respect: they would judge Inference sentences as “wrong” in the See Evidence trials.

All participants were at ceiling for the control sentences. The critical measure of performance on the Target sentences was not participants’ overall accuracy for each sentence type, but the *pattern* of responses across both Direct Perception and Inference sentences, particularly in the See Event and See Evidence conditions. We conducted a cluster analysis to identify response patterns using the *mclust* package in R (Scrucca et al. 2016). The input to the cluster analysis was each participant’s proportion of correct responses (according to the expected pattern described above) for Direct Perception and Inference sentences in the See Event and See Evidence trials only (four scores per participant), as these were the crucial conditions for assessing participants’ understanding of *see*. Data from adult and child participants were analyzed together to more easily identify children whose responses were similar to those of adults (and vice versa). The optimal model had five clusters of equal variance. The full set of participant data (all responses to Targets in all conditions) was then annotated with the cluster information, i.e. which cluster each participant belonged to as identified by the cluster analysis. The data were then analyzed with logistic mixed effects models using the *lme4* package in R (Bates et al. 2014) to evaluate the differences between clusters, that is, whether the clusters corresponded to significantly different ways of responding to the target sentences across the three conditions.

The majority of adult responses fit into two patterns. Nine of the adult participants gave responses consistent with our predicted pattern: they judged both target sentences as “right” in the See Event condition, and only Inference sentences as “right” in the See Evidence condition (Figure 4, Cluster 1). Three more adults gave responses that differed from the former group of adults only in one respect: they judged Inference sentences as “wrong” in the See Event condition significantly more often than the first group of adults ($\beta = -4.33$, $SE = 1.19$, $p < 0.05$; Figure 4, Cluster 2), suggesting they treated the two target sentences as mutually exclusive, a pattern consistent with adult responses in Experiment 1. As shown in Figure 4 (top row), the participants in these two clusters overwhelmingly judged the Inference sentences as “right” ($M = 0.82$) in the See Evidence condition, demonstrating an understanding that *see* with an embedded clause can report inference. Only one child clustered with the adults and gave responses that fit the predicted pattern for full knowledge of *see* complements.

All other child participants showed no understanding that the different complements corresponded to a difference in meaning. Their responses fit into three patterns. One third of children ($n = 7$), plus one adult, consistently judged *see that* as “wrong” in the See Evidence

¹ In that experiment, adults carried out a truth-value judgment task in which participants heard only one target sentence per trial; results showed that they judged both *see* and *see that* as correct for reporting direct perception.

condition ($M = 0.10$; Figure 4, Cluster 3), which was significantly more often than adults who fit the predicted pattern ($\beta = -4.66, SE = 0.76, p < 0.05$). This group otherwise made adult-like judgments of the Direct Perception and Inference sentences, indicating that their only difficulty was in understanding that “I saw that...” could be used to report inference.

Somewhat unexpectedly, about half of the children ($n = 12$) judged all of the target sentences in every condition as “right” (Figure 4, Cluster 4), even in the Doesn’t See condition where Mary saw nothing. However, these children were at ceiling for the control sentences, so their incorrect responses were not due to an overall “right” bias or a total failure to understand the task. When asked follow up questions at the end of the experiment, the children in this group confirmed that Mary was wearing her blindfold and could not see during the event in the See Evidence or Doesn’t See trials, so they were not confused or mistaken about her visual access in these trials. Instead, many of these children said that Mary’s *see* statement was correct because “I/we saw it” or because the described event *did* happen. Taken together, this indicates a ‘realist’ interpretation of the target sentences – that is, these children judged Mary’s *see* statements as “right” because the complement gave an accurate description of the event.

The remaining children ($n = 3$) and one adult were at chance for the Inference sentences in the See Evidence condition ($M = 0.54$), suggesting uncertainty about their meaning (Figure 4, Cluster 5). Children’s responses (across all clusters) were not predicted by age ($\beta = 0.05, SE = 0.18, p > 0.05$).

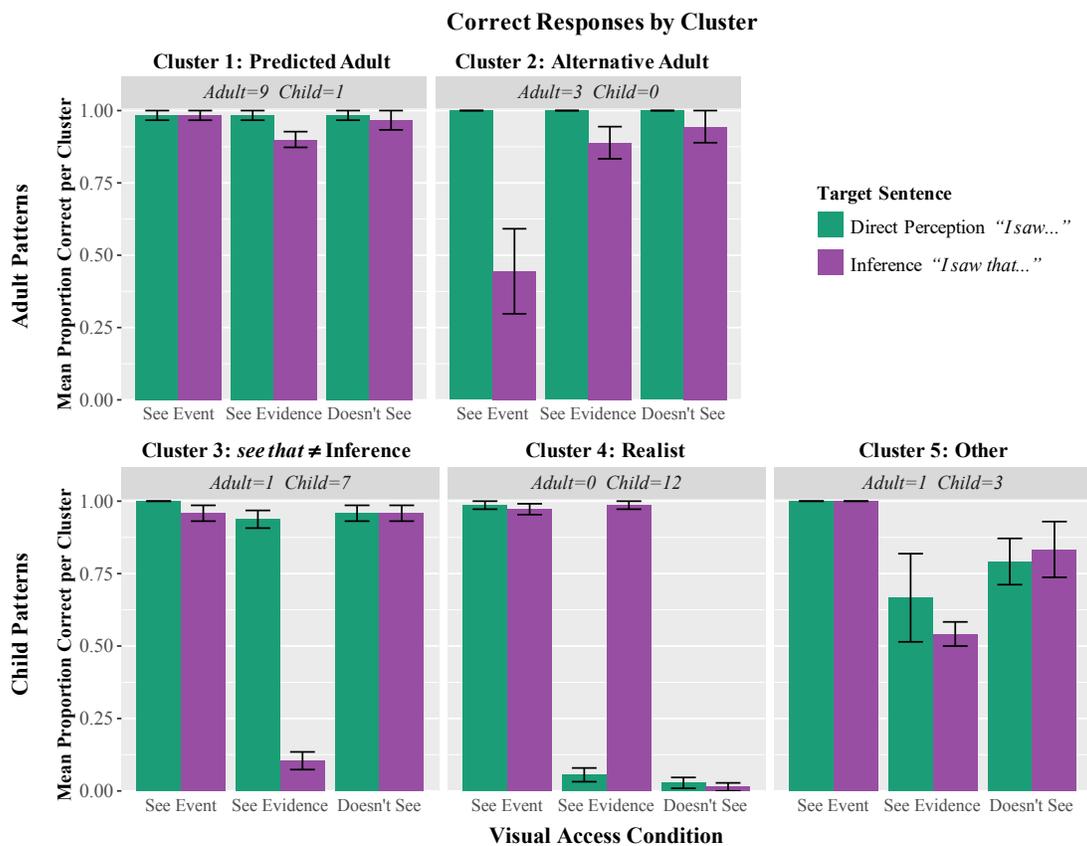


Figure 4. Mean proportions of correct responses for target sentences in Experiment 2, grouped by response pattern.

3.3. DISCUSSION. The results of Experiment 2 show that children under 7 years old do not demonstrate an awareness of the inference meaning of *see* + sentential complement despite linguistic and pragmatic conditions optimized to support this reading. That is, even when factors that could have biased children toward a direct perception interpretation of the target sentences were removed, most children did not show an understanding of the distinction between *see* with small clause and sentential complements. The majority of adults accepted *see that* sentences for reporting inference, but only one child did; the majority of children gave responses reflecting non-adult like interpretations of *see* and its complements. About a third of the children said *see* with a sentential complement was “wrong” in the See Evidence condition; this is consistent with the results of Experiment 1, in which children showed a tendency to interpret *see that* as having only a direct perception meaning, and suggests that linguistic knowledge, and not solely pragmatics, can account for children’s performance in that task. Additionally, about half of the children in Experiment 2 judged all *see* statements as “right” regardless of what the speaker had actually seen. In fact, 13 children commented during the task that the two target sentences were the same. Furthermore, age did not predict children’s response patterns, so it was not the case that, for example, only the youngest participants in this experiment were the ‘realists.’ Given that the children 7 and older in Experiment 1 performed like adults, the overall lack of an age effect for children ages 4-6 suggests that there is a qualitative change in children’s understanding of *see that* around the age of seven.

4. Conclusion. Our experiments show that 4-6-year-olds do not recognize that *see* can report visually-based inference when it takes a sentential complement (e.g. “I saw that someone peeled the banana”), even when pragmatic and contextual cues make inference interpretations highly salient. Our results indicate that adults and children over seven have both the direct perception and inference readings as part of the semantics of *see*, and make use of a variety of cues to select from these possible interpretations. In particular, adults and older children can use syntactic information (complement type) and pragmatic information (such as the implicature of the contrast of multiple frames and the perceptual experience shown in context) to determine when *see* means “perceive directly with the eyes” vs. “infer from (visual perception of) evidence.” Children between four and seven, however, are still learning the syntax and semantics of perception verbs like *see* and how distinct syntactic forms encode different kinds of perceptual experience. Our results suggest a significant change in children’s semantic representations around age seven, with earlier representations corresponding to *see* as encoding only direct visual perception, and later ones coming to include knowledge that *see* can report inference and an understanding of the relationship between a wider range of frames and their meanings.

One caveat is that we only tested children’s comprehension of statements describing another individual’s experience, and judgments about what another person would infer from visual evidence could be particularly difficult. Children might show an understanding of inferential *see* when asked about their own inferences rather than someone else’s, a possibility we are examining now. Even so, our results are consistent with cross-linguistic findings on children’s acquisition of evidential language, which has also shown a protracted developmental trajectory, particularly in the comprehension of forms used to report others’ perception or knowledge. Previous work has suggested that this difficulty in learning markers of direct perception vs. inference, regardless of marking type or language being acquired, may be the result of children’s difficulty integrating reasoning about other people’s experiences with their own linguistic knowledge (Ünal & Papafragou 2018, de Villiers et al. 2009). Thus, the potentially universal qualitative shift in

children's understanding of such language around age seven may reflect a change in their cognitive ability to synthesize linguistic and conceptual information about perception and knowledge.

References

- Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford University Press.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2014. Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software* 67(1). 1–48.
- Davis, E. Emory & Barbara Landau. 2020. Seeing and Believing: The Relationship between Perception and Mental Verbs in Acquisition. *Language Learning and Development*. 1–21. <https://doi.org/10.1080/15475441.2020.1862660>.
- de Villiers, Jill, Jay Garfield, Harper Gernet-Girard, Tom Roeper & Margaret Speas. 2009. Evidentials in Tibetan: Acquisition, Semantics, and Cognitive Development. (Ed.) S. A. Fitneva & T. Matsui. *Evidentiality: A window into language and cognitive development, New Directions for Child and Adolescent Development* 125. 29–47. <https://doi.org/10.1002/cd.248>.
- Elli, Giulia, Marina Bedny & Barbara Landau. 2021. How does a blind person see? Developmental change in applying visual verbs to agents with disabilities. *Cognition* 212. 104683. <https://doi.org/10.31234/osf.io/dpgx5>.
- Gertner, Yael, Cynthia Fisher & Julie Eisengart. 2006. Learning Words and Rules. *Psychological Science* 17(8). 684–692. <https://doi.org/10.1111/j.1467-9280.2006.01767.x>.
- Goldin-Meadow, Susan, Martin E.P. Seligman & Rochel Gelman. 1976. Language in the two-year old. *Cognition* 4(2). 189–202. [https://doi.org/10.1016/0010-0277\(76\)90004-4](https://doi.org/10.1016/0010-0277(76)90004-4).
- Landau, Barbara & Lila R. Gleitman. 1985. *Language and Experience: Evidence from the Blind Child*. Cambridge, MA: Harvard University Press.
- Papafragou, Anna, Peggy Li, Youngon Choi & Chung Hye Han. 2007. Evidentiality in language and cognition. *Cognition* 103. 253–299. <https://doi.org/10.1016/j.cognition.2006.04.001>.
- Pillow, Bradford H. 1999. Children's Understanding of Inferential Knowledge. *The Journal of Genetic Psychology* 160(4). 419–428. <https://doi.org/10.1080/00221329909595555>.
- Scrucca, Luca, Michael Fop, T. Brendan Murphy & Adrian E. Raftery. 2016. Mclust 5: Clustering, classification and density estimation using Gaussian finite mixture models. *R Journal* 8(1). 289–317. <https://doi.org/10.32614/rj-2016-021>.
- Sodian, Beate & Heinz Wimmer. 1987. Children's Understanding of Inference as a Source of Knowledge. *Child Development* 58(2). 424. <https://doi.org/10.2307/1130519>.
- Ünal, Ercenur & Anna Papafragou. 2016. Production-comprehension asymmetries and the acquisition of evidential morphology. *Journal of Memory and Language* 89. 179–199. <https://doi.org/10.1016/j.jml.2015.12.001>.
- Ünal, Ercenur & Anna Papafragou. 2018. Relations Between Language and Cognition: Evidentiality and Sources of Knowledge. *Topics in Cognitive Science*. 1–21. <https://doi.org/10.1111/tops.12355>.
- Ünal, Ercenur & Anna Papafragou. 2019. How Children Identify Events from Visual Experience. *Language Learning and Development*. Psychology Press 15(2). 138–156. <https://doi.org/10.1080/15475441.2018.1544075>.
- Winans, Lauren, Nina Hyams, Jessica Rett & Laura Kalin. 2015. Children's comprehension of syntactically-encoded evidentiality. *Proceedings of NELS* 45. 189–202.