**_Seeing_ vs. _Seeing That_: Children's Understanding of Direct Perception and Inference Reports**

Young children can reason about direct and indirect visual information (Ünal & Papafragou, 2019), but fully mapping this understanding to linguistic forms encoding the two knowledge sources appears to come later in development (Ünal & Papafragou, 2018; *inter alia*). 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 be used to report visually-based inference about an event. In two experiments, we sought to determine whether and when young English-speaking children, who already produce perception verbs like *see* in these frames, have linked the conceptual distinction between direct perception and inference to the different complements expressing this distinction.

In Experiment 1, we presented 36 children (4;0-9;01, *M*=6;5) with eight stories in which one character directly perceives an event, while a second encounters visual evidence that could lead to an inference about the event. After each story, participants heard both a Direct Perception sentence (e.g. "I saw Fido eating the cookies") and an Inference sentence (e.g. "I saw that Fido had eaten the cookies"), then were asked to identify which of the two characters said either the Direct Perception (4 trials) or Inference sentence (4 trials). Responses were marked correct if participants attributed the Direct Perception sentence to the direct perception character and the Inference sentence to the inferring character. Adult participants (n=6) always followed this pattern.

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 seven (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 seven and older (n=13, *M*=8;4) performed above chance (*M*=0.96, $t(12)$=17.725, $p$<0.05) and were no different from adults (*M*=0.96). A measure of joint performance on the two sentence types confirmed that the older children understood the distinction between *see* and *see that*, while the younger children did not ($\beta$=-0.97, $SE$=0.4, $p$<0.05; Figure 1).

In Experiment 2, we tested whether younger children would accept *see that* for reporting inference in a Truth-Value Judgment task. We showed 14 adults and 23 children (4;05-6;10; *M*=5;06) videos depicting an observer, Mary, watching a simple event in which an actor causes a visible change of state to an object (e.g. peeling a banana). Mary wears a blindfold at different points during the videos, varying by three within-subjects conditions. 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 always saw the full event. After every video, Mary made both a Direct Perception statement (e.g. "I saw someone peel the banana") and an Inference statement (e.g. "I saw that someone peeled the banana"), and participants were asked if she was right or wrong for each sentence. In the See Evidence condition, we expected that participants who understood the different meanings of the two frames would judge Direct Perception sentences as "wrong" and Inference sentences as "right" (correct), while participants who did not have this understanding would judge the Inference sentences as "wrong" (incorrect).

All participants were at ceiling in the See Event condition. In the See Evidence condition, adults overwhelmingly judged the Inference sentences as "right" (*M*=0.82), but only one child did so (Figure 2, 1st panel). All other children showed no understanding that the different complements corresponded to a difference in meaning, with children's responses fitting into

three patterns. One third of children consistently judged both *see* and *see that* as "wrong" in the See Evidence condition (*M*=0.10; Figure 2, 2nd panel). Half of the children judged all target sentences as "right" (Figure 2, 3rd panel), even in the Doesn't See condition where Mary saw nothing (*M*=0.02), but were at ceiling for control sentences, indicating a realist interpretation of the target sentences – that is, they judged Mary's *see* statements as "right" because the complement gave an accurate description of the event. The remaining children were at chance for the Inference sentences in the See Evidence condition (*M*=0.54), suggesting uncertainty about their meaning (Figure 2, 4th panel). Children's responses patterns were not predicted by age (*β*=0.05, *SE*=0.18, *p*>0.05).

These results demonstrate 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 syntactic and contextual cues make inference interpretations highly salient. Children under seven are still learning the syntax and semantics of perception verbs like *see* and how distinct syntactic forms encode different kinds of perceptual experience. The 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 frames and meanings. This is consistent with cross-linguistic findings on children's acquisition of evidential language, which has also shown a protracted developmental trajectory.

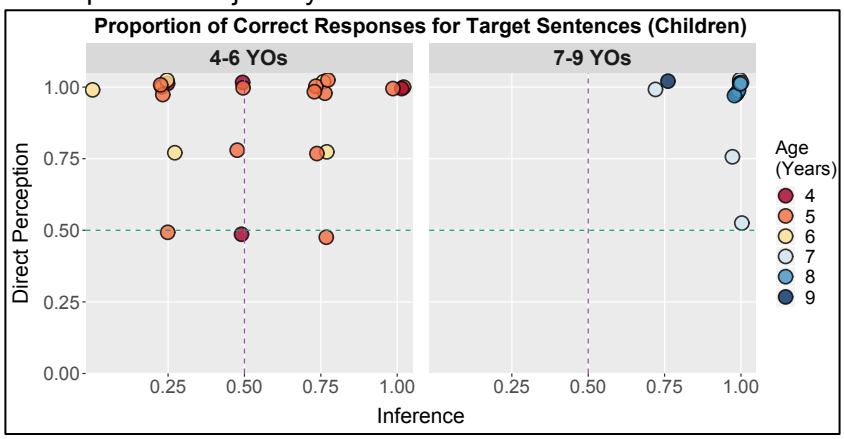


**Figure 1.** Proportion of correct responses for Direct Perception vs. Inference sentences in Experiment 1. Each point represents one participant; dotted lines show chance levels. Points in upper right corners indicate distinct interpretations for each sentence type; points in upper left corners indicate a direct perception bias.

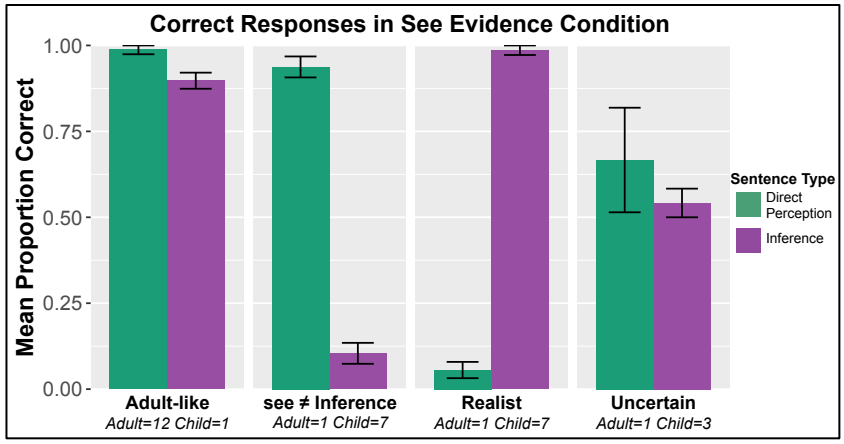


**Figure 2.** Mean proportions of correct responses for target sentences in the See Evidence condition in Experiment 2, grouped by response pattern. Correct = judging Direct Perception sentences as "wrong" and Inference sentences as "right" when Mary saw only evidence of the event.

## References


Ünal, E., & Papafragou, A. (2018). Relations Between Language and Cognition: Evidentiality and Sources of Knowledge. *Topics in Cognitive Science*, 1–21.

Ünal, E., & Papafragou, A. (2019). How Children Identify Events from Visual Experience. *Language Learning and Development*, *15*(2), 138–156.