Complement Size and Epistemic Effects in Perception Reports

Yitong Luo (yitong.luo@yale.edu)*, Clara Terlaak*, Xiaomeng Zhu* & Athulya Aravind
Department of Linguistics, Yale University

Abstract

Interpretation of perception reports with verbs like *see* varies depending on what the verb combines with. When *see* combines with a full tensed clause (e.g., *Ollie saw that a monkey was climbing a tree*), the perception report is thought to be epistemically *non-neutral*: the perceiver must form an event-consistent belief. But it is 'perceptually' non-committal — the perceiver need not have directly witnessed the event. When the verb combines with something smaller than a full tensed clause (e.g., *Ollie saw a monkey climbing a tree*), the pattern is claimed to flip. For the sentence to be true, the perceiver must have directly witnessed the event described by the complement clause. But, it does not commit them to a belief about what is perceived. Here, we experimentally test naive speakers' interpretations of these two types of perception reports by holding constant direct perception and manipulating epistemic neutrality. We found that irrespective of complement size, participants did not treat perception reports as epistemically neutral. We discuss two possible avenues to account for these unexpectedly strong readings, one semantic and one pragmatic. Ultimately, our findings suggest that the interpretive contrast between the two types of perception reports cannot be reduced to a simple perception-belief dichotomy.

Keywords: perception verbs; complement size; epistemic state; covered-box paradigm

Introduction

What information do we take in when we perceive the world, and how does perception change what we believe? For decades, philosophers and psychologists have treated perception and belief as tightly linked, asking how information delivered by the senses is transformed into changes in a perceiver's epistemic state (Dretske, 1981; Gopnik et al., 1994; Pylyshyn, 2003; Siegel, 2011; Slaughter & Gopnik, 1996; Song & Baillargeon, 2008). A central lesson from this work is that perception underdetermines belief. The same perceptual encounter can support very different belief updates. One may register an object's redness and roundness without recognizing it as an apple. Everyday experience is full of such partial or incomplete perceptual representations.

Natural language has a set of tools that lets us communicate about our perceptual experiences and resulting beliefs. Because these perceptual and mental states themselves lack perceptual correlates, language is a primary means of making them public. English has a rich class of perception verbs (*see, hear, look, perceive, sense*, among others) that can appear in a range of syntactic environments. Structural contrasts in linguistic form correlate with subtle distinctions in how and what we experience, and the mental state inferences that these experiences give rise to (Bedny et al., 2019; Gleitman, 1990; Landau & Gleitman, 1985).

We focus on two clause-embedding structures in which perception verbs such as *see* occur to form perception reports (Felser, 1998); see (1a)-(1b).

(1) a. Ollie saw [$_{CP}$ that a monkey was climbing a tree].

b. Ollie saw [$_{vP}$ a monkey climbing a tree].

In both, *see* takes a clausal complement that specifies the content of the perceptual experience. The crucial difference lies in the size of that complement. In (1a), the verb selects a full, finite clause, a CP: it includes the complementizer *that* and a tensed verb. In contrast, (1b) involves a smaller clause. Although it contains both a subject and a verb, it lacks tense or a complementizer. Building on prior work, we take this complement to be a *v*P (Felser, 1998; Moulton, 2009).[1]

Importantly, these differences in form correspond to differences in meaning (Barwise, 1981; Davis & Landau, 2020; Higginbotham, 1983; Moulton, 2009).[2] When the verb takes a CP complement, as in (1a), we derive an inference about the subject's epistemic state. Mere perception of an event is insufficient; the perceiver must also hold a corresponding belief about that event. If Ollie *saw that* a monkey was climbing a tree, then he must take what he saw to be a monkey climbing a tree. In this way, CP perception reports are epistemically non-neutral (Barwise, 1981). Sentences with vP-complements behave differently with respect to direct perception: with a sentence like in (1a), direct visual perception is not even necessary. For example, on finding the monkey already atop a tree, someone who never perceived the climbing event could still say (1a), based on an inference about what must have happened.

Perception reports with *v*P complements differ systematically in interpretation from their CP counterparts. In particular, sentences like (1b) can be used to report the perception of an event, without committing the perceiver to any corre-

*Equal contributions; ordering determined alphabetically.

[1]The precise analyses these authors propose vary somewhat, but we adopt a simplified version here as the finer-grained distinctions do not make a difference for our experiments.

[2]Though the structural contrasts generalize to other perception-verbs, the particular epistemic effects we describe here are specific to *see*.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

	see + *v*P	*see* + CP
Direct perception	✓	✗
Perceiver's belief	✗	✓
Epistemically neutral reading	available	unavailable

Table 1: Meaning components of *see* with *v*P vs. CP complement.

sponding belief about what was perceived. In other words, when *see* takes a *v*P complement, the resulting sentence is epistemically neutral.

The epistemic neutrality of such structures becomes especially clear when a follow-up sentence explicitly indicates the absence of such a belief. Consider the sentences in (2), from Barwise (1981). In (2a), the *v*P-complement sentence reports a perceptual event of a spy hiding a letter, even though the perceiver took what he was seeing to be a shoe-tying event. Crucially, this description does not sound contradictory, implying that the perception report does not require the perceiver to have a true belief about the content of the perceived event. In contrast, in (2b), the first clause uses a CP complement and *does* commit the perceiver to a belief about what they saw. When the follow-up sentence denies that belief, the result is a contradiction.

(2) a. Ralph saw [$_{vP}$ a spy hiding a letter under a rock], but thought she was tying her shoe.

b. *Ralph saw [$_{CP}$ **that** a spy **was** hiding a letter under a rock], but thought she was tying her shoe.

What the *v*P-variants *do* commit the perceiver to is having directly witnessed the event. Thus, (1b) would be judged false if Ollie only saw the monkey on top of the tree, but never saw it climbing that tree.

In short, CP and *v*P complements differ in how they relate perception and belief. CP complements are epistemically non-neutral: the perceiver is committed to a belief about what was perceived, but direct perception is not required, and inferences about the event are allowed. In contrast, *v*P complements do not require the perceiver to form a belief about the perceived event, but direct perception of the event itself is necessary. These patterns are summarized in Table 1.

Prior work has shown that adult English speakers are sensitive to the distinction between direct perception and inference in perception reports (Davis & Landau, 2021). However, the role of belief inferences, though central to formal semantic analyses (Moulton, 2009), has not, to our knowledge, been tested experimentally with naive adults. This is an important gap: impressionistically at least, epistemically non-neutral readings of *v*P complements appear quite common (Davis & Landau, 2021), and in everyday language they are often used interchangeably with CP complements.

The goal of the present study is to experimentally probe how the epistemic properties of *see* sentences vary depending on whether the verb takes a *v*P or a CP complement. To focus specifically on belief inferences, we hold the perceptual component constant: the perceiver is always a direct witness to an event. We then systematically vary how their perceptual experience affected their belief state. In this way, the evaluation of the *see*-*v*P/CP-sentences relative to the scenarios could be decided by the epistemic component alone. In two experiments, we ask: (1) whether *see* with a CP complement generates stronger belief inferences than *see* with a *v*P complement, and (2) whether these inferences are affected by the salience of the alternative structure. That is, does making the CP variant more accessible lead participants to interpret a *v*P-sentence as epistemically neutral, reasoning that if the sentence were epistemically non-neutral, the CP-variant would have been used.

Experiment 1

In Experiment 1, we use a covered-box paradigm to investigate the belief inferences triggered by perception reports involving *see*. Participants are presented with scenarios where a character witnesses an event but draws different epistemic conclusions – sometimes forming a full, accurate belief about what they saw, and other times forming only partial or no belief. Participants then judge whether a perception report with a CP or *v*P complement appropriately describes that perceiver, or whether it is more likely to describe a hidden individual. If *see*-CP structures require the perceiver to hold a belief about the event, we predict that in cases where the character derives only partial or no knowledge from that event, participants will favor the hidden individual as the one being described.

The sample size, procedures, and analysis plan for Experiment 1 were pre-registered.

Methods

Participants 84 native adult speakers of American English participated in Experiment 1. They were randomly assigned to two groups: 42 participants were assigned to the CP group (22 female, M_{AGE} = 43.69, SD_{AGE} = 15.47) and 42 to the *v*P group (22 female, M_{AGE} = 47.79, SD_{AGE} = 12.27). All participants indicated by self-report to have no history of language, hearing, or visual impairments. 10 additional adults were tested but excluded from data analysis in accordance with preregistered exclusion criteria: either failing comprehension checks on 'catch trials' probing story comprehension or failing on two or more of the filler trials. Participants were recruited from Prolific (www.prolific.com) and paid for their participation.

Materials, Design and Procedure We adopt a variation of the covered-box paradigm (Huang et al., 2013), which is designed to test whether a less accessible or less privileged reading of a sentence exists, even when it might otherwise be masked due to more favored or natural reading. In the classical covered-box task, participants are presented with two boxes, one whose contents are visible, and the other covered. Participants are then presented with a sentence and are asked to pick the box that is best described by the sentence. This

setup introduces a bias against the hidden box. Participants are expected to choose the open box with its visible content whenever it fits *any* reading of the sentence, including those readings that might be less natural or frequent. If, however, the content of the visible box does not support any reading of the sentence, participants are expected to, by exclusion, choose the covered box, even without knowing its content. This makes it a strong test: if the visible box could in principle support a given reading, but participants still choose the hidden box, it provides evidence that the reading is not in fact accessible to speakers.

Our experiment uses a similar logic. Participants are introduced to a puppet character Bubbles, who provides descriptions of his various friends. The participants' task is to decide whether Bubbles is describing a visible character, or a hidden one. For example, they may be asked to decide who 'Tommy' is based on Bubbles uttering: "My friend Tommy is wearing a red shirt". If the visible character is wearing a red shirt, participants should conclude that Tommy is that individual; if the visible character is wearing a green shirt, they should assume that Tommy is the hidden character.

All test trials describe events that take place when Bubbles' friends visit a zoo. Participants see scenarios involving two characters — one visible and whose mental states are known, and one hidden behind occlusion — and have to decide which character is best described by a statement Bubbles provides.

We constructed 12 scenarios along the schema in Fig.1. Each scenario began with an animal performing a specific activity, accompanied by a caption describing that event (A). A character (henceforth the *visible character*) then enters the scene (B) and expresses their belief state about the event. They may disclose (*i*) their full knowledge of what is happening (KNOWLEDGEABLE, "There's a monkey! The monkey is climbing a tree!"), (*ii*) their visual perception of the event without the corresponding knowledge about the nature of the event (PARTIAL, "There's a monkey! But what is it doing? It's too far away! What is the monkey doing? I don't know!"), or (*iii*) their total ignorance about the event (IGNORANT, "Oh, a tree! Is there anything else? I don't know!") (C). Subsequently, another character (henceforth the *hidden character*) is introduced; this character is hidden behind an occluder, and their belief state remains unknown to the participants throughout the frame (D). Finally, both the visible character and the character behind the occluder are shown on a single scene, alongside Bubbles, and participants are prompted to attend to the statement Bubbles will make about one of his friends (E). Finally, the target sentence, containing *see* and either a CP or *v*P complement, is presented in a speech bubble on the screen, and participants are instructed to select, by clicking on a corresponding radio button, whether the visible or the hidden character is being described by the sentence (F).

In addition to the test trials, participants also received 4 filler trials which crucially differed in that the scenarios and sentences did not involve perception reports. These trials showed the visible character engaged in an activity (e.g., flying a kite) and Bubbles' statement involved simple transitive verbs (e.g., *Lilly was flying a kite!*).

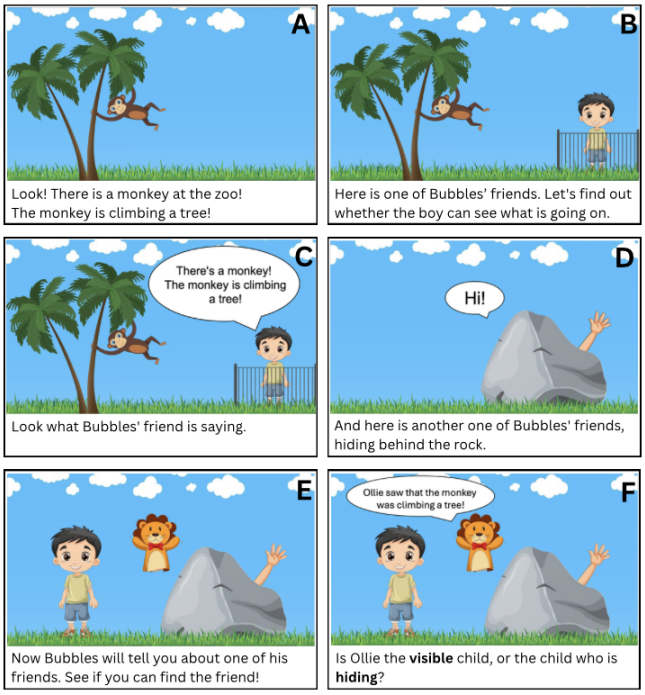

Figure 1: A sample test trial in Exp. 1

In addition to the knowledge state of the visible character (KNOWLEDGEABLE, PARTIAL, and IGNORANT) and complement type involved in Bubbles' perception report (CP vs. *v*P), we also manipulated sentence polarity (affirmative vs. negative) to control for the expected frequency of picking the visible character and the hidden character under every condition. Foreshadowing our predictions a bit, we expect that with affirmative target sentences, the visible character will always be picked under the KNOWLEDGEABLE condition, and the hidden character will always be picked under the IGNORANT condition. Negating an affirmative target sentence will always change the character the sentence fits. The predictions are laid out in more detail in the next section.

In sum, we manipulated three variables: Knowledge-state and polarity were within-subject variables, and complement type was between-subject. Table 2 exemplifies the target sentences (picture F in Fig.1).

Participants were randomly assigned to one of 6 lists, with different combinations of complement type, scenario, and condition. The order of trials within each list was randomized for each participant.

The study began with 2 practice trials, similar in structure to the fillers. Participants saw 4 further comprehension questions as attention checks; these questions queried something other than the belief state or perception of the story's characters in a given trial, e.g., which of two kinds of animal participated in the event perceived by the visible character.

Complement	Polarity	Target Sentence
CP	affirmative	Ollie saw that the monkey was climbing a tree.
CP	negative	Ollie didn't see that the monkey was climbing a tree.
*v*P	affirmative	Ollie saw the monkey climbing a tree.
*v*P	negative	Ollie didn't see the monkey climbing a tree.

Table 2: Target sentences by complement type and polarity.

Predictions

Conditions where the visible character is fully KNOWLEDGEABLE or fully IGNORANT serve as our baselines, where we don't expect to see differences between CP- and *v*P-sentences. In the KNOWLEDGEABLE condition, the visible character both directly perceives and forms a true belief about an event. Affirmative perception reports with both complement types would truthfully describe such a character. Negative sentences would be false of the visible character, so demand the selection of the hidden character. In the IGNORANT condition, the visible character neither perceives the event nor holds an event-consistent belief, and thus meets neither the epistemic requirement of CP-sentences nor the perception requirement of the *v*P-sentences. As such, both affirmative CP- and *v*P-sentences are false of the visible character, while negative ones are true.

The critical test case for our research question is the PARTIAL conditions, where the visible character perceives an event, but is ignorant about what exactly that event is about. Here, we expect participants' selection patterns to diverge depending on complement type. Affirmative CP-sentences should lead participants to select the hidden character, because the visible character fails to meet the epistemic requirement (i.e., they lack a belief consistent with the perceived event). Negative CP-sentences, on the other hand, would be true of that character. Affirmative *v*P-sentences, however, should allow for selecting the visible character. These structures are thought to be epistemically neutral, and only require direct perception of the event, and the visible character with PARTIAL knowledge meets this requirement. Conversely, negative *v*P-sentences should result in selecting the hidden character, as they deny visual perception of the event that did in fact take place for the visible character.

Altogether, we expect the PARTIAL condition to pattern together with the IGNORANT condition for CP complements, and with the KNOWLEDGEABLE condition for *v*P complements.

Results and Discussion

We analyze rates of choosing the hidden character. Results for Experiment 1 are shown in the top panel of Figure 2. Descriptively, when the visible character had full knowledge of the event (i.e., KNOWLEDGEABLE), participants overwhelmingly chose the visible character when the target sentence was affirmative, and the hidden character when the target sentence was negative. The pattern was the same for both CP and *v*P conditions. This is expected, since the epistemic inference of the (affirmative) CP-sentence is licensed in this scenario, and though the *v*P-sentence is epistemically neutral, it is compatible with the belief inference. When the visible character had no knowledge about the event (i.e., IGNORANT), participants showed the opposite pattern: choosing the hidden character given an affirmative sentence and the visible character given a negative sentence, and uniformly for both complement types. This, too, is in line with what we expect, as in this condition, neither the epistemic component nor the perception component holds true of the visible character. In the PARTIAL condition, participants selected the hidden character over 75% of the time in affirmative sentences, and less than 25% of the time with the negative sentences. Crucially, and surprisingly, this pattern of responses did not change across complement types. In other words, it seems as though for our participants, *v*P-sentences — like the CP-sentences — imposed an epistemic requirement, requiring that the perceiver form a belief consistent with the perceived event.[3]

To statistically analyze these patterns, we fit a Bayesian logistic mixed-effects model predicting selection of hidden character, with complement type, knowledge state (reference level being KNOWLEDGEABLE), and polarity (and their interactions) as fixed effects and participant as a random intercept. Fixed-effect predictors were evaluated using Type II Wald χ^2 tests. We found a significant two-way interaction between knowledge state and polarity ($\beta_{\text{PARTIAL}} = -12.15$, 95% CrI = $[-15.75, -9.44]$; $\beta_{\text{IGNORANT}} = -11.61$, 95% CrI = $[-15.18, -8.96]$), confirming that the effect of polarity on hidden character selection indeed differed by knowledge state. The visible character was found to render affirmative descriptions true only in the KNOWLEDGEABLE condition. The affirmative description was found not to describe — and the negative description truthfully describes — the visible character in the partial and ignorant conditions. Importantly, we did not find a three-way interaction among complement type, knowledge state, and polarity ($\beta_{\text{PARTIAL}} = -1.77$, 95% CrI = $[-6.16, 2.77]$; $\beta_{\text{IGNORANT}} = -0.59$, 95% CrI = $[-4.87, 3.92]$), which again confirms the uniform treatment of the two types of perception reports.

Participants in our experiment were sensitive to some, but not all, aspects of the structural differences between perception reports with *see*. They interpreted KNOWLEDGEABLE scenarios

[3] One reviewer helpfully pointed out that our *v*P-sentences are also compatible with an alternative structure where *the monkey* is an object NP modified by a relative clause: Ollie saw the monkey (that was) climbing the tree. While this is a possible parse, our results suggest that this is not the interpretation of affirmative *v*P sentences that participants adopted here. In particular, the interpretation under this alternative parse is compatible with the PARTIAL condition, in which case participants should have selected the visible character. However, we observe the opposite pattern. That said, further experiments that use a bare verb instead, thereby eliminating the modifier reading, would serve as a sanity-check.

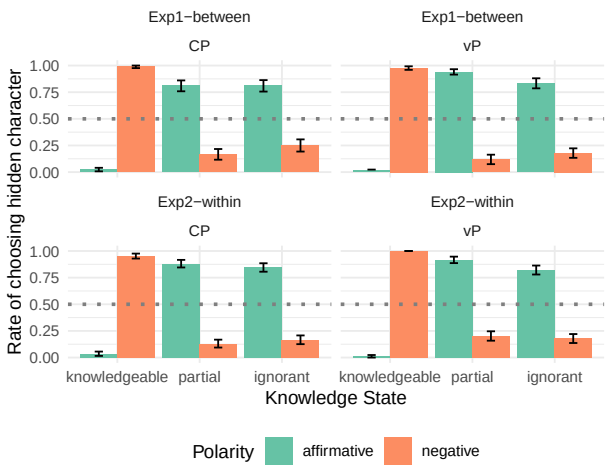

Figure 2: Rate of choosing the hidden character across Experiment 1 (top) and Experiment 2 (bottom) by knowledge state (x-axis), polarity (color) and complement type (CP on left vs. *v*P on right), and experiment (top vs. bottom). Chance is 50%. Error bars indicate ±1 standard error computed over participant means.

— where a perceiver directly witnesses an event and forms an accurate belief — as compatible with both CP and *v*P complements, and they interpreted IGNORANT scenarios — where a perceiver neither witnesses the event nor forms a belief — as compatible with neither. However, in the PARTIAL cases, participants did not distinguish between CP and *v*P complements, even though we expected these scenarios to be compatible with *v*P reports. In other words, participants in our study seemed to be interpreting *see*-*v*P-structures as also involving an epistemic component.

Do these results mean that *v*P-complements in perception reports carry a stronger, non-neutral meaning than previously reported? Before drawing that conclusion, we consider an alternative explanation. In our between-subjects design, participants who saw only *v*P sentences may have pragmatically strengthened these reports. After all, in everyday experience, perceiving an event often leads to forming a true belief, and without contrast to CP complements, the distinction between epistemically neutral and non-neutral readings may not have seemed relevant. It is possible that this effect would diminish if participants were presented with both CP and *v*P sentences in direct contrast, making the epistemic difference more salient. We test this possibility in Experiment 2, where complement type is manipulated within-subject.

Experiment 2

Sample size, procedures, and analysis plans for Experiment 2 were pre-registered.

Method

Participants An additional 84 native English-speaking adults were recruited from Prolific and randomly assigned to either the CP-first group (n = 42, 25 female, M_{AGE} = 48.26, SD_{AGE} = 15.07) or the *v*P-first group (n = 42 , 22 female, M_{AGE} = 45.18, SD_{AGE} = 11.73). 21 additional participants were excluded due to either the same exclusion criterion as in Exp 1 or having previously participated in Exp 1.

Materials, Design and Procedure All materials and procedures were identical to Experiment 1, except for the following. First, in addition to knowledge state and sentence polarity, complement type was also manipulated within-subjects. Participants were exposed to the different complement structures in blocks, and we varied whether a participant saw the CP block or the *v*P block first. Second, in order to raise participants' awareness of the varying syntactic forms of the target sentences, the introduction screen explicitly asked participants to pay attention to small differences between sentences. And finally, to further ensure that participants encoded the CP-*v*P difference, the target sentences were not only displayed in writing, but simultaneously presented auditorily.

If the lack of a complement type effect in Experiment 1 was due to pragmatic strengthening, caused by participants seeing only *v*P perception reports, then we expect to find block-order effects in Experiment 2. Specifically, participants who see *v*P-sentences in the second block, after already being exposed to the epistemically non-neutral CP-sentences, should access the epistemically neutral readings of these *v*P-sentences more often than participants who see *v*P-sentences first. In contrast, if *v*P-sentences inherently carry a stronger, non-neutral meaning, then we should not observe a difference between blocks. Behavior for CP-sentences is not expected to vary across blocks, since their epistemically non-neutral interpretation should not depend on a contrast with *v*P-sentences.

Results and Discussion

Results for Experiment 2 are shown in the bottom panel of Figure 2, which collapses across blocks. The results here are strikingly similar to Experiment 1. In the PARTIAL knowledge state, participants continued to select the hidden character with affirmative sentences and the visible character with negative sentences, irrespective of complement type.

We analyzed these data using a Bayesian logistic mixed-effects model with an identical structure to those in Experiment 1, except that block order was included as an additional interacting fixed effect. Consistent with Experiment 1, there is a robust two-way interaction between knowledge state and polarity ($\beta_{\text{PARTIAL}} = -8.45$, 95% CrI = $[-10.67, -6.46]$; $\beta_{\text{IGNORANT}} = -8.05$, 95% CrI = $[-10.27, -6.09]$). This interaction was further modulated by block, as indicated by the three-way interaction between knowledge state, polarity, and block ($\beta_{\text{PARTIAL}} = -7.18$, 95% CrI = $[-12.39, -2.66]$; $\beta_{\text{IGNORANT}} = -6.67$, 95% CrI = $[-11.75, -2.27]$). This interaction was driven solely by the fact that responses increased

in consistency in the second block. We found no indication that earlier exposure to the epistemically non-neutral structure shifts participants' behavior on the *v*P-sentences. In other words, there is no interaction with complement type within our model. Taken together, the results of Experiment 2 reinforce the findings from Experiment 1: even when participants encountered both CP and *v*P sentences within the same experiment, the structural contrast did not give rise to a corresponding difference in epistemic inferences.

General discussion

This study set out to test a claimed difference between two types of perception reports involving the verb *see*, which has been argued to give rise to subtle differences in what is inferred about a perceiver's mental state. It is often stated that perception reports with CP complements are epistemically non-neutral, and necessarily involve a belief on the part of the perceiver, whereas those involving *v*P complements are epistemically neutral (Barwise, 1981; Higginbotham, 1983; Moulton, 2009). However, across two experiments, we find that *v*P complements are also interpreted as epistemically non-neutral. Participants seemingly treat both *v*P and CP constructions as committing the perceiver to a belief about the content of the perceived event. These results call into question the standard characterization of the epistemic contrast between CP and *v*P perception reports.

At first glance, our findings might suggest that perception reports with *v*P complements have a stronger meaning than previously assumed, indeed, perhaps even stronger than that of their CP counterparts. On this view, *v*P reports would require not only direct perception of the event, but also that the perceptual episode results in a corresponding belief. While such an account would straightforwardly capture our experimental results, it raises two concerns.

First, it would undermine the tight link between CP complements and mental-state attribution that has been emphasized in prior work (e.g., Davis and Landau 2020, 2021) and that learners appear to be highly sensitive to (Gleitman & Papafragou, 2005; Harrigan et al., 2019). That said, the association between CPs and epistemic commitment is robust but not universal. Some attitude predicates take non-finite complements while still encoding an epistemic component. For example, the factive predicate *be surprised* remains factive, even when it takes a nonfinite (i.e., non-CP) complement, just like its CP counterpart. Thus, the sentence *Dana was surprised to pass the exam* entails that Dana passed the exam, and furthermore, requires the attitude holder to know this fact.

The second concern is more serious. A strong epistemic analysis of *see*–*v*P wouldn't account for the reported acceptability of sentences like (3), where the content of the event perceived by Ralph deviates from his belief.

(3) Ralph saw a spy hiding a letter under a rock, but thought she was tying her shoe. (Barwise, 1981, p. 374)

As we mentioned before, these observations have served to provide empirical support for a treatment of *v*P complements as epistemically neutral (Barwise, 1981). If speakers uniformly treated *see*–*v*P-structures as epistemically non-neutral, sentences like (3) should be incoherent. At present, however, these judgments have not been tested experimentally.

A second possibility is that the manipulation in Experiment 2 was not sufficient to reduce pragmatic strengthening, given how strong the connection between seeing and believing is. While linguistically, perception reports with *v*P complements may not be tied to belief, in many real-world contexts, humans treat perceiving as perceiving and forming a belief based on what was perceived. Visual perception serves as a primary input to conceptual representations and is closely associated with belief in its content (Dretske, 1981; Siegel, 2011). Even infants and toddlers use perceptual information to reason about mental states (Gopnik et al., 1994; Song & Baillargeon, 2008), and developmental work suggests that an understanding of perception precedes and scaffolds the development of theory of mind (Davis & Landau, 2020; Gopnik et al., 1994; Slaughter & Gopnik, 1996). In short, perception and belief are closely linked both conceptually and developmentally. This may lead speakers to default to an epistemically strengthened interpretation of perception reports even when such inferences are not grammatically required, reflecting a tight psychological link between perception and belief. This strengthening may, furthermore, be difficult to suppress.

If pragmatic strengthening is indeed responsible for the observed effect, then, it should in principle be possible to cancel the epistemically non-neutral inference (even if it is difficult). The results of Experiment 2 suggest that mere exposure to grammatically epistemically non-neutral structures is not sufficient to eliminate this strengthening. However, making the epistemically non-neutral interpretation incompatible with subsequent material in the sentence may succeed where implicit contrasts did not. In particular, a continuation of the sentence that is incompatible with the belief inference (as in (3)) should block, or at least substantially reduce, pragmatic strengthening, assuming speakers want to avoid contradictions. On the other hand, if perception reports with *v*P complements are grammatically associated with a belief component, then there would be no way out.

In ongoing work, we aim to tease apart these two possibilities by presenting participants with fuller sentences in which a continuation explicitly contradicts the epistemically non-neutral reading, as in (4).

(4) Ollie saw the monkey climbing a tree, but he did not realize it.

If the epistemically non-neutral reading of *v*P complements is a result of pragmatic strengthening, then this continuation should be able to cancel it, allowing the epistemically neutral reading to surface. If so, participants should now pick the visible character in the PARTIAL condition for affirmatives, reversing the pattern we observe in the current experiments. At present, the issue remains open. But one thing is for sure: the interpretive contrast between *v*P and CP perception reports cannot be reduced to a simple perception-belief dichotomy.

References

Barwise, J. (1981). Scenes and Other Situations. *The Journal of Philosophy*, *78*(7), 369–397. https://doi.org/10.2307/2026481

Bedny, M., Koster-Hale, J., Elli, G., Yazzolino, L., & Saxe, R. (2019). There's more to "sparkle" than meets the eye: Knowledge of vision and light verbs among congenitally blind and sighted individuals. *Cognition*, *189*, 105–115. https://doi.org/10.1016/j.cognition.2019.03.017

Davis, E. E., & Landau, B. (2020). Seeing and Believing: The Relationship between Perception and Mental Verbs in Acquisition. *Language Learning and Development*, *17*(1), 26–47. https://doi.org/10.1080/15475441.2020.1862660

Davis, E. E., & Landau, B. (2021). Seeing vs. Seeing That: Children's Understanding of Direct Perception and Inference Reports. *Experiments in Linguistic Meaning*, *1*, 125. https://doi.org/10.3765/elm.1.4864

Dretske, F. I. (1981). *Knowledge and the flow of information*. MIT Press. https://mitpress.mit.edu/9780262040631/knowledge-and-the-flow-of-information/

Felser, C. (1998). Perception and control: A Minimalist analysis of English direct perception complements. *Journal of Linguistics*, *34*(2), 351–385. https://doi.org/10.1017/S0022226798007075

Gleitman, L. R. (1990). The Structural Sources of Verb Meanings. *Language Acquisition*, *1*(1), 3–55. https://doi.org/10.1207/s15327817la0101_2

Gleitman, L. R., & Papafragou, A. (2005). Language and thought. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 633–661). Cambridge University Press.

Gopnik, A., Slaughter, V., & Meltzoff, A. (1994). Changing Your Views: How Understanding Visual Perception Can Lead to a New Theory of the Mind. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 157–181). Lawrence Erlbaum Associates, Inc. https://www.taylorfrancis.com/books/edit/10.4324/9781315804668/children-early-understanding-mind-charlie-lewis-peter-mitchell

Harrigan, K., Hacquard, V., & Lidz, J. (2019). Hope for syntactic bootstrapping. *Language*, *95*(4), 642–682. https://doi.org/10.1353/lan.2019.0069

Higginbotham, J. (1983). The Logic of Perceptual Reports: An Extensional Alternative to Situation Semantics. *The Journal of Philosophy*, *80*(2), 100–127. https://doi.org/10.2307/2026237

Huang, Y. T., Spelke, E., & Snedeker, J. (2013). What Exactly do Numbers Mean? *Language Learning and Development*, *9*(2), 105–129. https://doi.org/10.1080/15475441.2012.658731

Landau, B., & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child*. Harvard Univ. Press.

Moulton, K. (2009). Natural Selection and the Syntax of Clausal Complementation [Publisher: University of Mass Amherst]. https://doi.org/10.7275/H8XR-2A34

Pylyshyn, Z. W. (2003). *Seeing and Visualizing: It's Not What You Think*. The MIT Press. https://doi.org/10.7551/mitpress/6137.001.0001

Siegel, S. (2011). *The Contents of Visual Experience*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195305296.001.0001

Slaughter, V., & Gopnik, A. (1996). Conceptual Coherence in the Child's Theory of Mind: Training Children to Understand Belief. *Child Development*, *67*(6), 2967–2988. https://doi.org/10.2307/1131762

Song, H.-j., & Baillargeon, R. (2008). Infants' reasoning about others' false perceptions. *Developmental Psychology*, *44*(6), 1789–1795. https://doi.org/10.1037/a0013774