

**Please cite the Published Version**

Katsos, N, de Linares, BG, Ostashchenko, E  and Wilson, E (2023) Perspective-taking in deriving implicatures: The listener's perspective is important too. *Cognition*, 241. 105582 ISSN 0010-0277

**DOI:** <https://doi.org/10.1016/j.cognition.2023.105582>

**Publisher:** Elsevier

**Version:** Published Version

**Downloaded from:** <https://e-space.mmu.ac.uk/632755/>

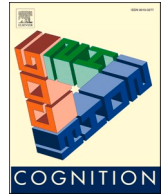
**Usage rights:**  [Creative Commons: Attribution 4.0](https://creativecommons.org/licenses/by/4.0/)

**Additional Information:** This is an open access article published in *Cognition*, by Elsevier.

**Data Access Statement:** Data available at: <https://osf.io/gnxjb/>.

**Enquiries:**

If you have questions about this document, contact [openresearch@mmu.ac.uk](mailto:openresearch@mmu.ac.uk). Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)



# Perspective-taking in deriving implicatures: The listener's perspective is important too

Napoleon Katsos<sup>a,\*</sup>, Blanche Gonzales de Linares<sup>a,b</sup>, Ekaterina Ostashchenko<sup>c</sup>,  
Elspeth Wilson<sup>a,b</sup>

<sup>a</sup> Department of Theoretical and Applied Linguistics, University of Cambridge, United Kingdom

<sup>b</sup> Faculty of Education, University of Cambridge, United Kingdom

<sup>c</sup> Department of Psychology Manchester Metropolitan University, United Kingdom

## ARTICLE INFO

### Keywords:

Pragmatics

Implicature

Perspective-taking

## ABSTRACT

Theories of meaning propose that listeners understand a speaker's implicit meaning thanks to mutually assumed norms of conversation that take into account what the speaker has said, as well as contextual factors, including what the speaker knows. Emerging psycholinguistic research shows that listeners derive a particular kind of implicit meaning, quantity implicatures, when their speaker is knowledgeable about the situation but tend to not derive it otherwise. In this article we focus on if and how listeners use the knowledge that is available only to themselves, i.e., the listener's perspective, while deriving implicatures. To do so, we explore the derivation of ad hoc quantity implicature in situations where the speaker does or does not have full knowledge, while, in the latter case, the listener has two types of privileged knowledge. Two versions of a study with neurotypical English-speaking adults show that listeners are influenced by their own perspective while deriving implicatures, depending on the type of knowledge available to them. We discuss the implications of these findings for models of pragmatic interpretative strategies.

## 1. Introduction

It is a common observation that much of the meaning that interlocutors exchange in conversation is not explicitly said. Grice's (1975) influential proposal argues that implicit meaning is communicated by the speaker and understood by the listener thanks to both of them assuming that they are cooperative and that they follow maxims of conversation, such as giving enough information for the purpose in hand (for recent reformulations of these maxims but nevertheless within the same broader spirit see also Carston, 1998; Frank & Goodman, 2012; Geurts, 2010; Horn, 1984; Levinson, 2000; Sperber & Wilson, 1995 among others).

To illustrate this, consider a situation in which there are several cards depicting fruit. Card (A) shows some apples, and card (B) shows some apples and some pears, while other cards depict other types of fruit but not apples. If your conversational partner asked you to 'pick up the card with apples and pears', you would be confident that they are asking for card B. If they asked you to 'pick up the card with apples', you could logically pick up either card A or card B on the grounds that they both

have apples. However, there is a good chance that you would pick card A, the only one with apples. Grice's (1975) communicative principles and maxims provide a rational reconstruction of this reasoning process, which relies on consideration of alternative ways that the speaker could have spoken, and expectations that speakers give the appropriate quantity of information. If the speaker meant to ask for card B in the situation above, they should have said 'pick up the card with apples and pears' which is an informative way to refer to it in this situation. Because the speaker did not use a description which would be informative for card B, you can conclude that the speaker who asks for 'the card with apples' also means 'the card with apples and nothing else on it', an inference which is known as an ad hoc quantity implicature. This inference would then lead you to pick card A.

Importantly, Grice's account factors in several aspects of the extra-linguistic context of the conversation, including what the speaker knows. Imagine that the speaker does not know that card A, the one with only apples, exists, because, for example, it has fallen off the table. If the speaker were to say 'give me the card with apples' in this situation, then one would give them card B, on the grounds that it is the only card that

\* Corresponding author.

E-mail address: [nk248@cam.ac.uk](mailto:nk248@cam.ac.uk) (N. Katsos).

<https://doi.org/10.1016/j.cognition.2023.105582>

Received 31 October 2022; Received in revised form 26 July 2023; Accepted 28 July 2023

Available online 6 September 2023

0010-0277/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

has apples, regardless of the fact that it also has pears. Therefore, the speaker's knowledge of the situational context is important, both in order for the listener to derive quantity implicatures when the speaker is fully knowledgeable about the situation, and in order not to derive them when the speaker lacks important knowledge.

While the Gricean approach to implicature was meant as a philosophical reconstruction of how a listener would arrive to an implicature (see [Geurts & Rubio-Fernández, 2015](#); [Katsos, 2008](#)), it seems to hold well as a psycholinguistic model too. A small but growing body of research that we review in Section 1.1. reports that listeners mostly do adopt the speaker's perspective when interpreting utterances with implicature (indicatively, see [Bergen & Grodner, 2012](#); [Breheny, Ferguson, & Katsos, 2013](#); [Papafragou, Friedberg, & Cohen, 2018](#)). Before we review these studies though, in the next section we turn to a larger set of studies on reference and perspective-taking,<sup>1</sup> which reports that while listeners overwhelmingly take the speaker's perspective, there are discernible influences from the listener's own perspective. This research on reference highlights a gap in our current understanding, because the majority of the studies on implicature and perspective-taking to date have not investigated how the listener uses any privileged information they have, and how they integrate it with the speaker's perspective.

## 2. Language interpretation and perspective-taking

The role of speaker's and listener's knowledge states and what they assume to be 'common ground' ([Stalnaker, 1978](#); see also [Clark, 1996](#); [Allan, 2013](#), for a survey of various conceptualisations) has been investigated extensively in reference resolution. Here, a large body of work has produced a mixed set of outcomes and a debate on how the perspectives of the interlocutors are used. Influential work by [Keysar, Barr, Balin, and Brauner \(2000\)](#) tracked participants' eye movements while they looked at a grid containing objects placed between them and a confederate. Some of the slots in the grid were hidden from the speaker's point of view, but not from the participant's point of view. In the critical condition, listeners were given an instruction by the speaker (e.g., 'pick up the small candle') that could refer to an object in common view (the smaller of two candles in common view). However, for the listener there was a better match with an object in her privileged view (an even smaller candle). Participants looked significantly more and longer at the small candle in the privileged view, which shows that, at least in some cases, listeners are egocentric, in that they do not take the speaker's perspective into account in the first stage of the processing of the utterance. Several other studies have also argued that listeners interpret language from their own perspective, at least in the early stages of online process and sometimes at the end of the interpretative process too (e.g., by actually reaching out to select an object that the speaker could not see; [Epley, Morewedge and Keysar, 2004](#); [Lin, Keysar, & Epley, 2010](#); [Keysar, Lin, & Barr, 2003](#); [Apperly et al., 2010](#); among others).

Other studies, however, report that when everything else is equal in the experimental paradigm, the speaker's perspective is consistently prioritised over the listener's own. For example, [Hanna, Tanenhaus, and Trueswell \(2003\)](#) argued that the original results from the [Keysar et al. \(2000\)](#) study were not due to participants being egocentric, but rather to

the fact that the privileged object was always a better perceptual match to the instructions than the object in the listeners privileged view. To demonstrate this, they ran an experiment with a similar design, but where the common ground and privileged ground objects were both identical red triangles. When hearing instructions such as "put the blue circle above the red triangle", participants were always more likely to look at the common ground red triangle and faster to choose it. Likewise, [Heller, Grodner, and Tanenhaus \(2008\)](#) provide further evidence that listeners can use the speaker's perspective efficiently and from the earliest stages of processing. There are several ways in which the conflicting data can be reconciled ([Barr, 2008](#); [Brown-Schmidt & Hanna, 2011](#); [Hanna et al., 2003](#); [Heller, Parisien, & Stevenson, 2016](#)) to which we will return in General Discussion.

A conclusion from perspective-taking and reference assignment is that while listeners predominantly take the speaker's perspective into account, privileged information from the listener's perspective can also affect utterance interpretation. Increasing the strength of the cues that perspective-taking is needed (e.g., see [Brown-Schmidt & Hanna, 2011](#)) or the motivation that participants have to do so (see [Bezuidenhout, 2013](#)), as well as minimising listeners' extraneous demands on executive control and memory (e.g., see [Brown-Schmidt, 2009](#); [Ryskin, Brown-Schmidt, Tullis, & Benjamin, 2015](#); [Wardlow, 2013](#)) are some of the ways that can decrease the effect of the listener's perspective. Speakers and listeners have also been shown to monitor the division of labor between them as an interaction unfolds, with speaker's producing more informative expressions or listener's increasingly adopting the speaker's perspective according to each other's behaviour and the conversational situation (see [Hawkins, Gweon, & Goodman, 2021](#)). Importantly, these studies focus on referring expressions such as adjective-noun combinations ('the big/red triangle') and ambiguous nouns ('tape' as in measuring tape or cassette). While pragmatics and expectations of informativeness are important for reference assignment in some of these paradigms (e.g. the ones using adjectives contrastively such as [Heller et al., 2016](#)), there is no need for a the listener to derive an implicature per se.

### 2.1. Implicature and perspective-taking

Perspective-taking and quantity implicature derivation is increasingly becoming a topic of interest. The focus has been on whether the listener factors in the speaker's knowledge or lack of it in deriving implicatures. In the first study on this topic, [Bergen and Grodner \(2012\)](#) adapted an experimental design from [Breheny, Katsos, and Williams \(2006\)](#) and presented participants with a context in which the speaker was either fully knowledgeable (for example they say 'I meticulously compiled the investment report') or partially knowledgeable ('I skimmed the investment report'). This was followed by a sentence containing 'some' (e.g., 'Some of the real-estate investments lost money'). This sentence could lead to a quantity implicature where an utterance with 'some' implies that the more informative alternative 'all' is not the case, leading to interpreting 'some' as implying 'not all'. These sentences were followed by one whose subject NP was highly accessible as a referent if the implicature had been derived (e.g., 'The rest' [of the real-estate investments]...). [Grodner and Bergen](#) reported longer reading times for the implicature trigger ('some of the') and its continuations in the full knowledge condition, compared to the less knowledgeable condition. Moreover, reading times on the anaphor 'the rest' were shorter in the knowledgeable condition, confirming that an implicature was derived in the previous sentence because the implicature makes the complement set of the reference, those investments that were not successful, salient. The opposite pattern was found in the partial knowledge condition. The potential trigger segment was read faster and the continuation sentence slower, which suggests the absence of an implicature. These findings suggest that listeners derive quantity implicatures when they know that or can assume that the speaker is knowledgeable but do not derive them otherwise.

<sup>1</sup> We use the term 'perspective-taking' to refer to the ability to attribute knowledge to an interlocutor via reasoning about what is visible to them or not. 'Theory of Mind', 'mind-reading' and other terms are used in the literature typically with a wider meaning to encompass the attribution of beliefs and intentions to other people via any means. 'Perspective taking' as used in this article is a subset of these which concerns the attribution of knowledge to an interlocutor based on the understanding that seeing leads to knowing. The type of perspective taking that is required in the experiments reported in this article is Level-1, which assesses what someone can or cannot see. Other levels of perspective-taking are discernible, e.g., assessing how what someone else sees differs ([Moll & Meltzoff, 2011](#)) – but they are not relevant for this article.

Using a different paradigm, Goodman and Stuhlmüller (2013) presented participants with scenarios where the speaker declared that he had either full or partial knowledge of the situation, before making a statement that could potentially be interpreted with a quantity implicature. For example the speaker would say that they have opened and looked at 2 out of 3 letters (partial knowledge) or 3 out of 3 letters (full knowledge) before saying that ‘Some of them have...’X’ inside them’. Participants were asked to place a bet on the number of letters that had ‘X’ inside. In the full knowledge condition participants placed significantly more bets on 2 out of 3 letters compared to 3 out of 3 letters, while in the partial knowledge condition no difference was found, which is evidence that in absence of full knowledge of the situation, listeners suspended the derivation of implicatures (though the extent to which they suspend the inference varies, see Dieuleveut, Chemla, & Spector, 2019). Further research has investigated the time-course of listeners’ taking the speaker’s perspective into account (Breheny et al., 2013, using eye-tracking in the visual world paradigm), the neural signatures using ERPs (Spychalska, Reimer, Schumacher, & Werning, 2021), as well as investigating whether young children (Barner, Hochstein, Rubenson, & Bale, 2018; Hochstein, Bale, Fox and Barner, 2016; Kampa & Papafragou, 2020; Papafragou et al., 2018; Wilson, Lawrence, & Katsos, 2022;) and neurodiverse populations (Hochstein, Bale, & Barner, 2018) incorporate perspective-taking in implicature derivation.

Importantly, Bergen and Grodner (2012), Goodman and Stuhlmüller (2013), and other studies (e.g., Barner et al., 2018; Hochstein et al., 2018) focus on the speaker’s knowledge of the situation (whether it is full or partial) and whether it affects the listener’s interpretation. The role of the listener’s own knowledge is not critical because there is nothing in the listener’s perspective that differs from the speaker’s perspective, and which could influence the derivation of implicature. Consider, however, a situation where the listener knows more than the speaker, and where this additional knowledge would allow an implicature interpretation of the critical sentence (from the listener’s perspective). Only a handful of studies have investigated how the listener uses any privileged information they have, and how they integrate it with the speaker’s knowledge. Breheny et al. (2013) reported evidence for the seamless integration of the speaker’s perspective with implicature derivation in adults by using eye tracking. Participants in the experiment saw events unfolding in a video and listened to descriptions of the events by a confederate speaker, while their eye movements were monitored. In critical trials an utterance such as “The woman put a spoon in the box” leads to the ad hoc quantity implicature “The woman put a spoon in the box and nothing else” in the context of two boxes, one with a spoon inside and another with a spoon and a fork. The listener and the speaker watch events of a person putting different kinds of objects into containers one at a time, unfolding over time. In some trials, a screen is placed in front of the speaker so that they do not see the last act of putting an item in a container. In the critical condition testing whether the listener considers the speaker’s perspective, either the speaker had seen all of the events in the video and they knew exactly what was put in each container; or they had seen up to the point where two boxes were identical (e.g., both had a spoon inside), and not had not seen the event to its conclusion (e.g., whereby a fork was added to the other box, such that one box had a spoon but the other eventually had a spoon and a fork). The listener, however, always had full knowledge. The prediction was that in this condition, the listener would take the lack of speaker knowledge into account and not derive the implicature when the speaker says ‘the woman put a spoon in the box’, which would lead to no anticipatory glances to any of the two boxes. This prediction was fulfilled, despite the fact that in this experimental paradigm the listener did have knowledge (the fact that one of the boxes contained a spoon and a fork) which could have led the listener to derive an implicature (from her own perspective).

Another paradigm that creates a mismatch between the speaker’s perspective and listener’s perspective is deployed by Kampa and Papafragou (2020) who present participants with two displays, each of which

have two parts. In the critical trials one of the two parts of one display holds an object (e.g., a spoon) and the other part another (e.g., a bowl). There are two such displays each with the image of the same confederate positioned behind the display, as if they sat opposite the participant. In one of the two displays, the confederate can see all that the listener can see, and in the other display a part of it with an item is occluded from the confederate (e.g., the participant can see the spoon but they cannot see the bowl). The participants hear one of two utterances, and they are asked to say which of the two displays the confederate is talking about, e.g., when they say ‘I see a spoon’ vs ‘I see a spoon and a bowl’. (see also others, e.g., Papafragou et al., 2018 for a similar design). Kampa and Papafragou report that children and adults are at ceiling at considering the speaker’s perspective, and select the speaker who cannot see the bowl as the likely speaker of the first utterance (‘I see a spoon’) while selecting the speaker who can see both the bowl and the spoon for the latter utterance, even though from the participant’s perspective there is always a spoon and a bowl in both displays. These findings again support the view that children and adults engage with perspective-taking in this task and their behavior is not affected by what is visible in their own privileged view.

However, an indication that the listener’s perspective might be important too comes from a recent study by Wilson et al. (2022) which investigated children’s ability to adopt the speaker’s perspective in conditions which resembled those of Breheny et al. (2013). In the critical case, the listener could see four cards, two of which shared the same type of object (e.g., pears on one card vs pears and bananas on the other). However, the card with pears only was exclusive to the listener’s perspective and not visible to the speaker. In this condition, when the speaker asked for ‘the card with pears’ a listener who relies on their own perspective would erroneously derive an ad hoc quantity implicature and point to the card with only pears, whereas a listener who took the speaker’s perspective would not derive an implicature and point to the card with pears and bananas. The study was designed to investigate children’s pragmatic development and only a small group of adults were recruited as a control group ( $n = 18$  in Experiment 2). While the adults’ accuracy was at rates of over 85% as a group, it did not reach ceiling as it did on all other conditions in the task, as there were occasions where adult participants selected the card that exclusive to their own perspective. Looked at in terms of individuals, 4 out of the 18 adults failed to meet the criterion of ‘passers’ in this task (accuracy of 5/6 or 6/6 trials). With the caveat that this was a small group of participants, this could be considered as first tentative evidence that adults may be affected from what is known to them from the listener’s perspective. Even more recently, other ongoing research by Jarvinen, O’Shea, and Barr (2022) also suggests that listeners are affected by privileged knowledge in their own perspective when it comes to deriving contrastive implicatures from adjectives.

But what would implicatures without perspective-taking look like? From the point of view of Gricean and neo-Gricean theory the question does not even arise. In order to derive an implicature it is presupposed that the listener is adopting the speaker’s perspective when interpreting their utterances. However, more recent accounts do allow for implicatures to take place on occasion (or even primarily) without considering the speaker’s perspective. The details of these accounts vary substantially. Some of these accounts argue that representing the speaker’s perspective and integrating it with the listener’s is not routinely done at the psychological level because (a) it is cognitively costly and/or (b) for many communicative situations, the interlocutors co-presence, familiarity and history of successful interactions, among other factors, more or less safely warrant the assumption that the listener’s own perspective is identical to the speaker’s perspective, without requiring the listener to actively consider the speaker’s perspective in its own right (see Breheny, 2006; Jary, 2013; Katsos & Andrés-Roqueta, 2021; Kissine, 2016; Moore, 2018; Sperber, 1994; Sperber et al., 2010).

Yet other accounts tie the generation of an implicature closer to the form of language used and less upon inter-subjective factors such as the

speaker's knowledge. Indicatively, default accounts of implicature (Levinson, 2000) assume that listeners derive some implicatures (including quantity, scalar implicatures) upon the parsing of the corresponding form of words (e.g., 'some'). Assumptions about cooperation or what knowledge is shared between interlocutors have a role only as secondary filters, occasionally suspending an implicature which was generated in the first place. Similarly, grammar-oriented accounts, place the core mechanism of implicature derivation in the grammar itself (syntax and/or semantics), postulating the existence of a silent grammatical operator, similar to 'only', which generates and negates alternatives (indicatively, Chierchia, 2004; Chierchia, Fox, & Spector, 2012). Like default accounts, grammatical accounts do not deny that contextual assumptions and shared knowledge are important, though exactly what role they play is not fully specified.

What is common to all the accounts above is that implicatures may arise without the listener truly considering the speaker's perspective in a communicative situation. In practical terms, in referential communication paradigms, this would mean that listeners assume that everything that is visible to themselves is available as a possible referent of the speaker's instructions. Similar distinctions between alternative interpretative strategies that do or do not use perspective-taking or other forms of Theory of Mind are possible for other pragmatic phenomena besides quantity implicature (see Del Sette, Bambini, Bischetti, & Lecce, 2020, and Lecce, Ronchi, Del Sette, Bischetti and Bambini, 2019, for metaphor; see also Deliens, Antoniou, Clin, & Kissine, 2017, and Deliens, Antoniou, Clin, Ostashchenko, & Kissine, 2018, for sarcasm and irony; among others).

## 2.2. Research aims and hypotheses

In a set of experiments that we report here, we adapted the referential communication paradigm that has been employed in much of the research on reference resolution, to add an implicature component to it, and created two conditions where the influence of the listener's perspective may become evident. In one of these conditions, the interpretation one would reach from the listener's perspective only is different from the one they would reach if they took the speaker's perspective (as in Breheny et al., 2013, and Wilson et al., 2022); and in the second condition the interpretation leads to ambiguity from the listener's perspective which can be resolved once the speaker's perspective is taken into account. We added this novel condition to explore if the way that the interpretations arrived at from the listener's and the speaker's perspective interact is also important (i.e., whether it matters if one of the interpretations leads to ambiguity or to clash with the other interpretation).

We predict that if listeners rely exclusively on the speaker's knowledge when deciding on whether to interpret the speaker's utterance with an implicature or not, then what the listeners see in their own privileged ground will not be relevant to their choice of referent (and the difference of whether what they see from their own perspective leads to ambiguity or a different choice than from the speaker's will also be irrelevant). On the other hand, if listeners are also affected from their own perspective, then we would expect to see somewhat fewer pragmatically appropriate interpretations when there is important information in the listener's perspective. We might have even more specific predictions depending on whether the instructions from the listener's perspective are ambiguous or whether they would lead to a different interpretation. Tentatively, we would expect that accuracy will be higher when the listener's perspective leads to ambiguity which can be resolved once the listener adopts the speaker's perspective, compared to when the listener's perspective leads to an unambiguous choice which contrasts with the choice that would be made from the speaker's perspective. We return to this topic in the General Discussion.

As regards the first prediction, that what is in the listener's perspective is not important, this is compatible with the original philosophical account put forward by Grice. However, when it comes to the

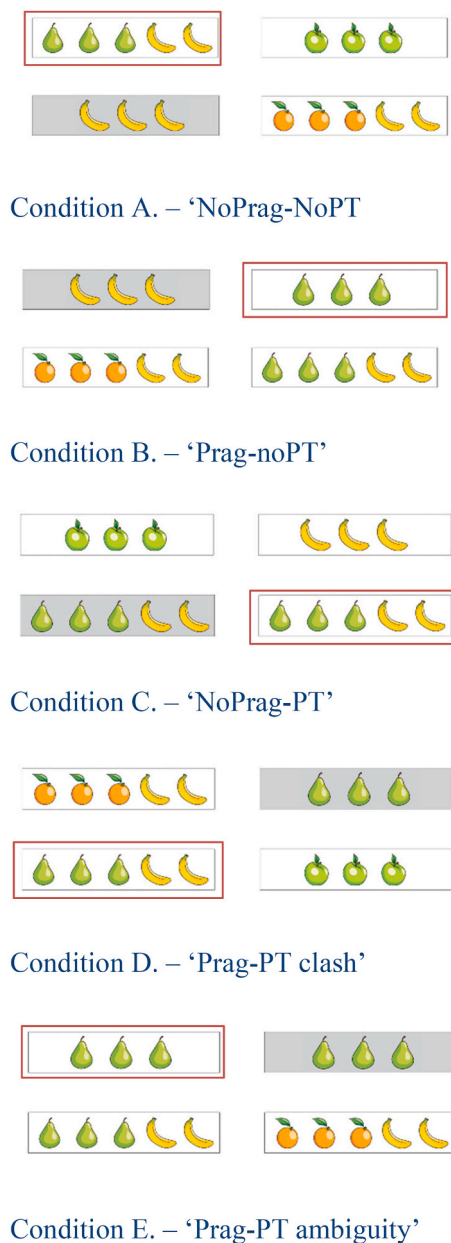
second prediction, as we noted earlier, there are views of pragmatic processing broadly in line with the Gricean linguistic-philosophical view on implicature which make additional assumptions about how and when the components of reasoning involved in implicature are activated in language processing. These views consider that listeners may use their own (egocentric) perspective or the speaker's (allocentric) perspective as part of their strategies for interpreting utterances pragmatically, depending on a number of factors. Listener-internal factors may be developmental age (see Breheny, 2006) or neurodiverse cognitive profiles (see Happé, 1993; Andrés-Roqueta & Katsos, 2020), or the availability of cognitive resources such as memory or executive control which regulate perspective-taking (Ryskin et al., 2015; Wardlow, 2013); in these cases a processing strategy based on the speaker's perspective may not be available or it may be too costly in terms of resources (see also Sperber et al., 2010) and a listener may apply pragmatic maxims such as quantity of information to an utterance under the assumption that the speaker knows exactly what she herself knows. In addition, there are also external and situation-dependent factors whereby a listener has more than one interpretative strategies available to herself but employs one where she considers that the speaker knows whatever the listener knows (effectively, an egocentric strategy) as a matter of course on grounds of cognitive economy, unless there are reasons to use a different one (see Katsos & Andrés-Roqueta, 2021; Kissine, 2016; Sperber, 1994; Sperber et al., 2010 – see also Breheny, 2006; Jary, 2013; Moore, 2018; on pragmatics without mind-reading). This diverse set of views, which we group together here because they stipulate that listeners have a range of interpretative strategies at their disposal, is compatible with the second predictions we made above, namely that the listener's own perspective does affect their interpretation, in at least some occasions and/or for at least some listeners. In the next section we report two experiments that put these predictions to the test.

## 3. Experiments 1A and 1B

We ran two similar studies, Experiment 1A and 1B, with the main difference between them that the stimuli in Experiment 1B were presented auditorily rather than in writing. We report these studies as two versions of the same experiment. The experiments were a computer-based version of the referential communication task widely used in studies on perspective-taking and the interpretation of adjectives and ambiguous nouns (Apperly et al., 2010; Hanna et al., 2003; Heller et al., 2008; Heller et al., 2016; Keysar et al., 2000; Keysar et al., 2003). It was complemented by expressions that may trigger an ad hoc quantity implicature which could resolve referential ambiguity in some trials. An earlier version of the task was used by Wilson, Lawrence, & Katsos (2023) who investigated children's ability to use the speaker's knowledge to derive or not ad hoc quantity implicatures. Compared to Wilson et al., in our experiments we added a new critical condition with privileged knowledge for the listener and we changed another condition to help rule alternative explanations of participant choices. The rationale for these changes is mentioned in the Discussion section below. We also made modifications that make the task more ecologically valid for adult participants.

### 3.1. Method

In this task the participant is presented with displays with four cards, and can see what is depicted in all four cards in front of her, whereas there is only one card that the speaker cannot see which is indicated by the shaded area in this part of the display (see Fig. 1). The speaker is asking the participant to pick a card. In some conditions the correct response to the speaker's instructions requires pragmatic inferencing, in some trials it requires taking the perspective of the speaker, and in some it requires both. In a baseline condition neither implicature nor perspective-taking is needed. We can illustrate this with reference to Fig. 1 below, where the speaker's request is 'Pick the card with pears' in



**Fig. 1.** A set with one display for each condition for the instruction “Pick the card with pears”. The pragmatically appropriate selection is highlighted in a red square. Shaded areas indicate that this card is not visible to the speaker. The short names of each condition reflect the interpretative strategies needed for successful selection (‘PT’ = perspective-taking; ‘Prag’ = ad hoc quantity implicature), if any. ‘clash’ and ‘ambiguity’ note the outcome of the interpretative process from the listener’s perspective if they derive an implicature but do not adopt the speaker’s perspective. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

all the conditions for ease of exposition. In the figure, red borders are added around the target card for each condition to show which response is coded as correct.

In condition A, which we call ‘NoPrag-NoPT’ only one card features the item mentioned in the instruction. This is a condition which requires no implicature or perspective-taking, because simple lexical semantics suffice for selecting the correct card. In condition B, which we call ‘Prag-NoPT’, two cards semantically match the instruction and are visible to both speaker and listener. Deriving an ad hoc implicature leads to the unambiguous selection of a card, and no perspective-taking is required.

The display in condition C, which we call ‘NoPrag-PT’, shows two identical cards that semantically match the instruction. However one of these two cards is in privileged ground for the listener, as shown by the shaded area indicating that this card is not visible to the speaker. This condition requires perspective-taking but no implicature. In condition D, which we call ‘Prag-PT clash’, two cards match the instruction semantically. From the speaker’s perspective, the target card is the one with pears and bananas because it is the only card they can see. However, from the listener’s perspective, if they derive an implicature, the target card is the one with pears only (importantly, the implicature interpretation is not available if the listener were to adopt the speaker’s perspective). In this condition therefore, the cards selected from the speaker’s and the listener’s perspective clash. Finally, in condition E, which we call ‘Prag-PT ambiguity’ three cards match the instruction semantically. Listeners can select a single target card if they derive an ad hoc implicature and consider speaker’s perspective as well. A difference between condition D, ‘Prag-PT clash’ and condition E ‘Prag-PT-ambiguity’ is in the outcome of interpreting the utterance from the listener’s perspective only. In the former the listener’s perspective selects a different target card than what would have been selected the speaker’s perspective, while in the latter the listener’s perspective leads to ambiguity as regards to which card to select.

### 3.1.1. Materials

Eight sets of the 5 conditions were used as stimuli material in this study resulting in a total of 40 trials per participant. One-half of the sets depicted fruit (pears, oranges, bananas, and apples) and the other depicted animals (pigs, cats, ducks, and hens), with each of these items being the target item in the instructions for each set. The positions of shaded and target cards were counterbalanced across trials. In Exp. 1B only another eight displays were added as a measure of attentional control where participants were asked to pick the card that cannot be seen by the speaker.

### 3.1.2. Procedure for Exp. 1A & 1B

The experiments were created on Gorilla TM Experiment Builder (Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2020). Participants completed them online from a location of their own choice. They were asked to provide information about themselves on age, gender, languages spoken from birth, and whether they were diagnosed with or they thought they might have a neurological or neurodevelopmental condition, and they were given information about the experiment and use of their data. If they gave approval, they were then familiarised with the items that they would be shown and their labels (e. g., ‘pears’, ‘apples’) as well as the four quadrants display, each with one card inside a quadrant which depicted some items. Participants were told they would read (Exp 1A) or hear (Exp 1B) instructions from a speaker who would ask the participant to give her one of the cards. One of the four quadrants would always be shaded to indicate that it cannot be seen by the speaker. To ensure that the participants understood that the card in the shaded part of the display is not visible to the speaker, they were shown how the display would look from the speaker’s perspective with the shaded area (and no card). They were also asked if a person in the position of the speaker would be able to see what is on the card behind the shaded area. They then completed a practice block consisting of 4 trials. Once participants confirmed their understanding of the task by accurately completing practice trials, they advanced to the main task. Participants were instructed to begin each trial by using the computer mouse to click on the ‘Next’ button located at the centre of the screen. The ‘Next’ button disappeared when clicked and the screen with four cards and the instruction immediately appeared (in writing, Exp 1A or auditorily, Exp 1B) until the participant completed their response by clicking on a card, upon which the ‘Next’ button would return to its original position. Forty test trials were presented in a different random order for each participant. Participants were instructed to respond as accurately as possible every time. Participants clicks on the cards were

recorded as well as the overall time that lapsed between the pressing of a 'Next' button to the selection of a card, which corresponds to the time when participants see a display and read or hear the instructions.

### 3.1.3. Procedure modifications for Exp 1B

Five modifications were made compared to Exp 1A. First, we used auditory instructions (prerecorded by a female adult native speaker of English) which is a better approximation of a real-life conversation. Second, in order to dispel any concerns that participants who select a card that the speaker cannot see might be doing so on the grounds that it might be one that the speaker might appreciate having, participants were explicitly told that the speaker could not ask for a card she cannot see. Moreover, eight displays were added as measure of attentional control where participants were asked to pick the card that cannot be seen by the speaker. Finally, to explore whether any effects of implicature and/or perspective-taking might be captured in reaction times in addition to a focus on accuracy that was emphasised in Exp. 1A, we asked participants to pick cards as accurately and as fast as possible. Finally, following participants' feedback in short debriefing interviews after Exp 1A, we introduced different sets of images for two items, to avoid potential issues with disambiguating pictures of ducks and hens.

## 3.2. Participants

**Experiment 1A:** Sixty participants, native speakers of English were recruited through [Prolific Academic Ltd \(2014\)](#). Data from one participant were excluded from the analysis because some of their responses were given in under 150 ms which is considered too fast for having meaningfully considered a trial. The final sample consisted of 59 participants (31 female), aged 18 to 54 years ( $M = 30;6$ ).

**Experiment 1B:** Fifty-two participants, native speakers of English were recruited through [Prolific Academic Ltd \(2014\)](#). Data from 3 participants were excluded from the analysis because some of their responses were given in under 150 ms which is too fast for having meaningfully considered a trial. The final sample consisted of 49 participants (24 female), aged 18 to 61 years ( $M = 38$ ).

No participants had a clinical history of neurological or neurodevelopmental condition (nor did they self-identify as having one) according to self-report.

### 3.2.1. Ethics

Ethics approval was given by the Ethics Committee of the Faculty of Modern and Medieval Languages of the University of Cambridge.

### 3.2.2. Analyses

Accuracy and reaction times to the selected items were recorded. Response times (RTs) were calculated over correct trials only. We separately analysed the effects of the experimental factors on accuracy and reaction times. For accuracy, a binary-dependent variable on the single trial level, we applied linear mixed-effects models using the `glmer()` function of the `lme4` package () implemented in the R environment ([R Core Team, 2021](#)). Models were compared based on  $\chi^2$ , z-values, and p-values. Multiple comparisons were adjusted using Tukey's Multiple Contrasts in R package `emmeans` for post-hoc testing ([Lenth, 2016](#)). We fitted a linear mixed-effects regression model to our data, with accuracy and reaction time as the dependent variables and the fixed effect for condition (A-E). Single term deletion was used to determine the significance of random effects in the model ([Bates, Mächler, Bolker, & Walker, 2015](#)). All models failed to converge when allowing for random intercept and slope for these effects, so random slopes were removed. Model comparisons on the maximal models that converged ([Barr, Levy, Scheepers, & Tily, 2013](#)) were performed to explore whether the inclusion of the fixed factor made a significant contribution to the model.

## 3.3. Results

Accuracy for Exp. 1A and 1B. is presented in [Table 1](#). and [Fig. 2](#) below.

We then applied the `glmer()` function to analyse the effect of condition on accuracy on the single trial level.

**Experiment 1A:** This analysis revealed a significant main effect of condition ( $\chi^2(4) = 53.08$ ,  $p < 0.001$ ). Post-hoc comparisons showed that accuracy in condition D - 'Prag-PT clash' was significantly lower than in all other conditions (see [Table 2.](#)). Participants were also somewhat more accurate in the baseline condition A - 'NoPrag-NoPT' compared to condition C - 'NoPrag-PT' and condition E - 'Prag-PT ambiguity'. Accuracy was also higher in condition B - 'Prag-NoPT' compared to condition C - 'NoPrag-PT', but no differences were detected between other conditions. Furthermore, erroneous choices were scrutinized in condition D - 'Prag-PT clash', the only condition with a sizeable number of errors (over 10%). Out of 67 incorrect trials, 63 (94%) involved the selection of the card in the shaded area.

**Experiment 1B:** The same analysis revealed a significant main effect of condition ( $\chi^2(4) = 67.23$ ,  $p < 0.001$ ). Post-hoc comparisons showed again that participants were less accurate in condition D - 'Prag-PT clash' compared to all other conditions (see [Table 3.](#)). In this version of the experiment they were more accurate in the baseline Condition A - 'NoPrag-NoPT' compared to all other conditions (see [Table 3](#)). Again, erroneous choices were scrutinized in the only condition (D 'Prag-PT clash') with over 10% errors. Out of 100 incorrect trials, 92 involved the selection of the card in the shaded area.

The `lmer()` function was applied to explore the effect of condition on log-transformed reaction times.

**Experiment 1A:** The model including condition as a fixed factor did not improve the fit of a model including random effects only ( $\chi^2(4) = 2.78$ ,  $p = 0.59$ ), indicating that latency to identify the referent was not significantly different across conditions. We also analysed the results after removing excessively long responses (calculated as those 2.5 standard deviations above the mean of the total experiment RT), on the grounds that these might reflect outlier responses based on processes other than the ones relevant for success in the task. For Exp. 1A this led to the removal of 55 out of 1765 correct trials (those whose RT exceeded 4945 ms; mean RT for Exp 1A = 2138 ms, SD = 1123). Again, a model including condition as a fixed factor did not improve the fit of a model including random effects only ( $\chi^2(4) = 1.47$ ,  $p = 0.83$ ).

**Experiment 1B:** Inclusion of condition did not lead to a significantly better model fit compared to the model without this predictor included ( $\chi^2(4) = 5.13$ ,  $p = 0.27$ ). We also analysed the results after removing excessively long responses calculated similarly to Exp. 1A. This led to the removal of 41 out of 1386 correct trials (those whose RT exceeded 3242

**Table 1**

Mean accuracy by trial for Experiments 1A and 1B by condition.

| Experiment | Condition               | Mean | Standard Deviation | Standard Error of the Mean |
|------------|-------------------------|------|--------------------|----------------------------|
| Exp. 1A    | A - 'NoPrag-NoPT'       | 0.98 | 0.12               | 0.007                      |
|            | B - 'Prag-NoPT'         | 0.96 | 0.19               | 0.011                      |
|            | C - 'NoPrag-PT'         | 0.90 | 0.29               | 0.016                      |
|            | D - 'Prag-PT clash'     | 0.78 | 0.41               | 0.023                      |
|            | E - 'Prag-PT ambiguity' | 0.93 | 0.25               | 0.014                      |
| Exp. 1B    | A - 'NoPrag-NoPT'       | 0.98 | 0.14               | 0.007                      |
|            | B - 'Prag-NoPT'         | 0.93 | 0.24               | 0.012                      |
|            | C - 'NoPrag-PT'         | 0.93 | 0.25               | 0.012                      |
|            | D - 'Prag-PT clash'     | 0.75 | 0.43               | 0.021                      |
|            | E - 'Prag-PT ambiguity' | 0.92 | 0.26               | 0.013                      |

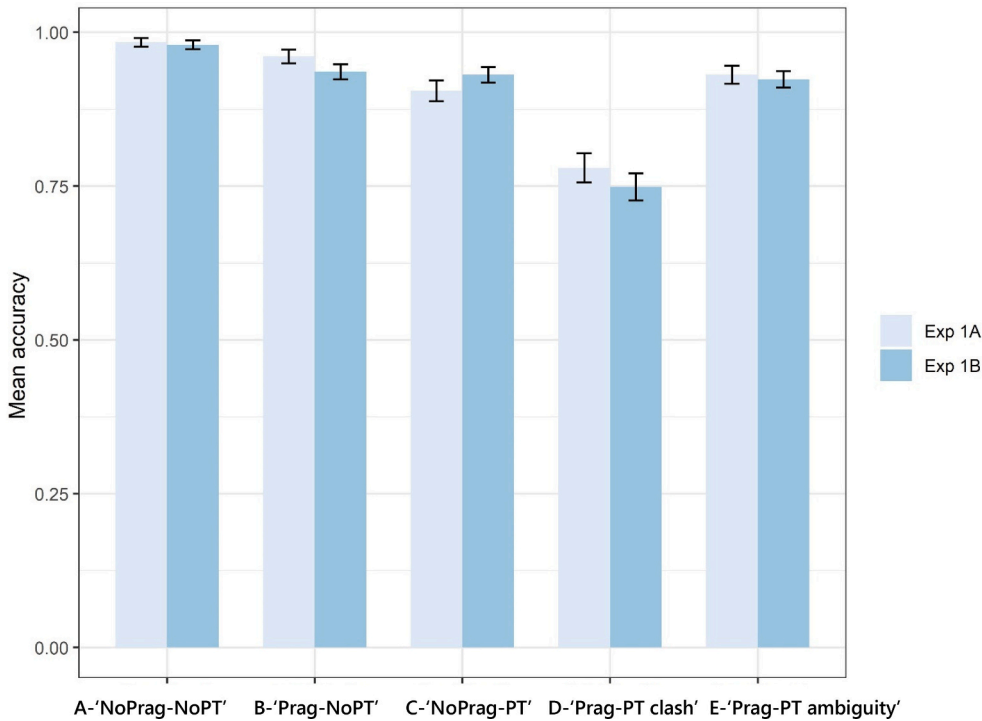


Fig. 2. Mean accuracy for Experiments 1A and 1B by condition; error bars indicate standard error of the mean.

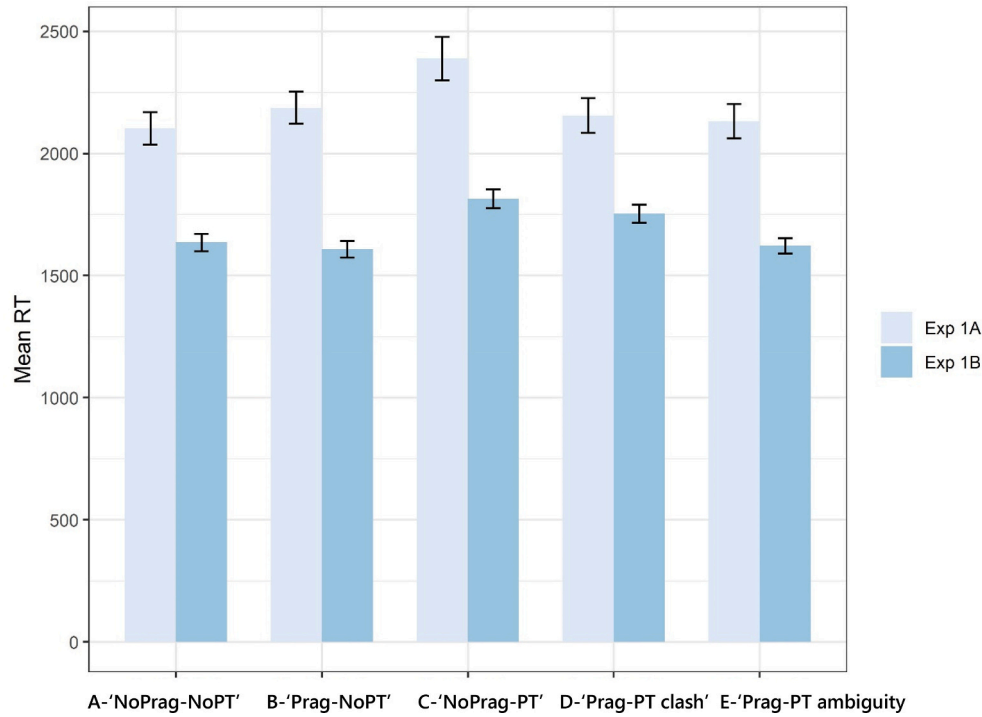


Fig. 3. Mean reaction times in Experiments 1A and 1B by condition in milliseconds.

ms; mean RT for Exp 1B = 1655 ms, SD = 635). Again, a model including condition as fixed factor did not improve the fit of a model including random effects only ( $\chi^2(4) = 5.58, p = 0.23$ ).

3.4. Discussion for Experiment 1A and 1B

As expected, participants performed near ceiling on conditions A – ‘NoPrag-NoPT’ and B – ‘Prag-NoPT’, the former a baseline where no

perspective-taking or implicature is required, the latter a condition where an ad hoc quantity implicature is required but no perspective-taking. They also performed well in the condition that required perspective taking without implicature, C – ‘NoPrag-PT’. In all these cases accuracy was at or above 90% in both Experiments, indicating that participants could derive ad hoc implicatures and that they could take the speaker’s perspective, in conditions where one or the other but not both were required. The conditions of main interest were D – ‘Prag-PT

**Table 2**

Post-hoc analyses output for the model fitting accuracy on the single-trial level in Exp 1A.

| Post hoc comparisons – Conditions                                 | $\beta$ | SE   | z-ratio | p value |
|---|---------|------|---------|---------|
| Condition A - 'NoPrag-NoPT' – Condition B - 'Prag-NoPT'           | 1.02    | 0.57 | 1.78    | 0.3826  |
| Condition A - 'NoPrag-NoPT' – Condition C - 'NoPrag-PT'           | 2.20    | 0.52 | 4.12    | 0.0003  |
| Condition A - 'NoPrag-NoPT' – Condition D - 'Prag-PT clash'       | 3.59    | 0.52 | 6.93    | <0.0001 |
| Condition A - 'NoPrag-NoPT' – Condition E - 'Prag-PT ambiguity'   | 1.74    | 0.54 | 3.23    | 0.0106  |
| Condition B - 'Prag-NoPT' – Condition C - 'NoPrag-PT'             | 1.18    | 0.40 | 2.99    | 0.0227  |
| Condition B - 'Prag-NoPT' – Condition D - 'Prag-PT clash'         | 2.28    | 0.37 | 6.80    | <0.0001 |
| Condition B - 'Prag-NoPT' – Condition E - 'Prag-PT ambiguity'     | 0.72    | 0.41 | 1.76    | 0.3927  |
| Condition C - 'NoPrag-PT' – Condition D - 'Prag-PT clash'         | 1.39    | 0.29 | 4.82    | <0.0001 |
| Condition C - 'NoPrag-PT' – Condition E - 'Prag-PT ambiguity'     | -0.46   | 0.34 | -1.34   | 0.6632  |
| Condition D - 'Prag-PT clash' – Condition E - 'Prag-PT ambiguity' | -1.83   | 0.32 | -5.83   | <0.0001 |

**Table 3**

Post-hoc analyses output for the model fitting accuracy on the single-trial level in Exp 1B.

| Post hoc comparisons – Conditions                                 | $\beta$ | SE   | z-ratio | p value |
|---|---------|------|---------|---------|
| Condition A - 'NoPrag-NoPT' – Condition B - 'Prag-NoPT'           | 1.37    | 0.44 | 3.11    | 0.0160  |
| Condition A - 'NoPrag-NoPT' – Condition C - 'NoPrag-PT'           | 1.45    | 0.44 | 3.34    | 0.0075  |
| Condition A - 'NoPrag-NoPT' – Condition D - 'Prag-PT clash'       | 3.58    | 0.42 | 8.61    | <0.0001 |
| Condition A - 'NoPrag-NoPT' – Condition E - 'Prag-PT ambiguity'   | 1.59    | 0.43 | 3.69    | 0.0020  |
| Condition B - 'Prag-NoPT' – Condition C - 'NoPrag-PT'             | 0.09    | 0.32 | 0.28    | 0.9987  |
| Condition B - 'Prag-NoPT' – Condition D - 'Prag-PT clash'         | 2.22    | 0.29 | 7.77    | <0.0001 |
| Condition B - 'Prag-NoPT' – Condition E - 'Prag-PT ambiguity'     | 0.23    | 0.31 | 0.73    | 0.9506  |
| Condition C - 'NoPrag-PT' – Condition D - 'Prag-PT clash'         | 2.13    | 0.28 | 7.62    | <0.0001 |
| Condition C - 'NoPrag-PT' – Condition E - 'Prag-PT ambiguity'     | 0.14    | 0.31 | 0.45    | 0.9915  |
| Condition D - 'Prag-PT clash' – Condition E - 'Prag-PT ambiguity' | -1.99   | 0.27 | -7.34   | <0.0001 |

clash' and E - 'Prag-PT ambiguity', where both perspective-taking and implicature are possible but the selective application of only one them would lead the participant to different referential outcomes.

A main finding is that condition D - 'Prag-PT clash' had significantly lower accuracy rates compared to all the other conditions. This includes lower accuracy than B - 'Prag-NoPT', and C - 'NoPrag-PT', which were constructed to be direct comparison points on implicature and perspective-taking on their own, respectively. Condition D - 'Prag-PT clash' was also lower than condition E - 'Prag-PT ambiguity'. These differences were significant in both Exp. 1A and 1B. However, condition E - 'Prag-PT ambiguity' was not different from B - 'Prag-NoPT', and C - 'NoPrag-PT'. These results indicate that what is in the listener's perspective does play a role in deriving quantity implicatures. If it did not, both conditions D - 'Prag-PT clash' and E - 'Prag-PT ambiguity' ought to have similar rates of accuracy as conditions B - 'Prag-NoPT' and C - 'NoPrag-PT'. This was not borne out, with condition D - 'Prag-PT clash' being consistently the condition with lowest accuracy.

Importantly, the fact that condition E - 'Prag-PT ambiguity' was not

**Table 4**

Reaction time (in milliseconds) by condition in Experiments 1A and 1B. Next, we turn to reaction times (see Table 4, and Fig. 3, below).

| Experiment    | Condition               | Mean (ms) | Standard deviation | n   | Standard error of the mean |
|---------------|-------------------------|-----------|--------------------|-----|----------------------------|
| Experiment 1A | A - 'NoPrag-NoPT'       | 2095      | 1153               | 299 | 66                         |
|               | B - 'Prag-NoPT'         | 2185      | 1127               | 292 | 66                         |
|               | C - 'NoPrag-PT'         | 2262      | 1247               | 275 | 75                         |
|               | D - 'Prag-PT clash'     | 1990      | 766                | 237 | 49                         |
|               | E - 'Prag-PT ambiguity' | 2134      | 1199               | 283 | 71                         |
| Experiment 1B | A - 'NoPrag-NoPT'       | 1604      | 655                | 383 | 33                         |
|               | B - 'Prag-NoPT'         | 1604      | 625                | 365 | 32                         |
|               | C - 'NoPrag-PT'         | 1769      | 669                | 363 | 35                         |
|               | D - 'Prag-PT clash'     | 1700      | 642                | 292 | 37                         |
|               | E - 'Prag-PT ambiguity' | 1606      | 564                | 362 | 29                         |

significantly different in accuracy than condition B - 'Prag-NoPT' or C - 'NoPrag-PT' in either study, and that it was higher than D - 'Prag-PT clash' in both studies, suggests the presence of an additional factor, besides the role of the listener's perspective. Tentatively, the difference between condition D - 'Prag-PT clash' and E - 'Prag-PT ambiguity' might be about whether the listener's perspective suffices for successfully achieving the task in hand, which is to choose a single card. Recall, that in condition E - 'Prag-PT ambiguity' if the speaker's perspective is not considered, the instructions are ambiguous from the listener's perspective, even after the listener derives an implicature, since there are two cards with the named entity, e.g., with just pears, on them. Faced with no clear choice between two plausible cards, some participants might simply guess (which, when correct, would bring their apparent accuracy closer to that of condition C - 'NoPrag-PT'). Others might use the lack of an unambiguous choice as a metacognitive cue that their interpretation is deficient, reasoning that in every other trial there is a single correct card that matches the instructions, and therefore there must be just one here too. This could prompt participants to adopt the hitherto ignored speaker's perspective, thereby choosing the correct card. Again, this would bring their accuracy close to or equal to that of condition C - 'NoPrag-PT', which is what we observed.

In contrast, in condition D - 'Prag-PT clash' if the speaker's perspective is not considered, the listener does have a card that fits the target instructions if they derive an implicature. With a plausible card that they can select (from their own perspective), the listener does not need to guess nor do they have a cue to prompt them to revise their interpretation. This could explain why in condition D - 'Prag-PT clash' we observed the lowest accuracy throughout conditions, in both experiments. This explanation invokes metacognitive processes, which involve an individual's insights into and control over their own mental processes (Flavell, 1979), including processes that help control, monitor, and regulate strategies to meet task demands and goals (see Tarricone, 2011) such as the predominant goal of these experiments which is to select a single card in each trial.

If this understanding is on the right track, it suggests that the listener's perspective can have a role in derivation of quantity implicatures, and that metacognitive monitoring can moderate this effect by occasionally prompting listeners to change interpretative strategies. See also the General Discussion for an additional hypothesis on the difference between conditions D - 'Prag-PT clash' and E - 'Prag-PT ambiguity'.

Finally, the reaction times collected did not reveal any significant

differences among conditions. Rather than concluding that there is no effect of implicature and/or perspective taking on processing, which would be against a large body of research on these topics (indicatively, [Bott & Noveck, 2004](#); [Breheny et al., 2006](#); [Huang & Snedeker, 2009](#); on implicature; [Keysar et al., 2000](#), [Keysar et al., 2003](#) on perspective-taking in referential communication), we modestly conclude that the online administration of the experiments was not ideal for capturing such effects, at least not when reaction times were recorded at the top-most level, the whole duration of the picture-presentation and instructions trial, rather than on segments of it. Further research, most likely under laboratory conditions and capturing segment-by-segment temporal data, should be in better position to explore online effects.

### 3.4.1. Alternative interpretation for accuracy rates

One overall concern about the interpretation of the results might be that participants in Exp. 1A and 1B are not engaging in perspective-taking, but rather in a shallow strategy which consists in globally ignoring the card in the shaded area throughout the experiments' trials. Nevertheless, we believe it is unlikely that this is what is happening, for several reasons: First, recall that we added 8 trials per participant within Exp. 1B where the participants are asked to 'pick the card that the speaker can't see'. One motivation for this condition was to make sure that participants were paying attention to which card the speaker could not ask for. Let us call this condition F - 'Attentional control'. Here, participants were at ceiling accuracy (3 errors out of 392 trials; 99% accuracy). However, in addition to ensuring that participants are paying attention to the experiment's critical manipulation, the inclusion of condition F in Exp. 1B minimises the effectiveness of a hypothetical strategy of globally ignoring the card in the shaded area, because participants are frequently required to interact with this card. Still, let's suppose that participants were applying a strategy of ignoring the shaded card not globally, across the experiment, but in all trials bar condition F - 'Attentional control'. In this case, we ought to see this manifest as some kind of suppression/inhibition of the card in the shaded area, which ought to be reflected in reaction times required to pick this card. However, average reaction time for this condition was 1124 ms (sd = 443), which is faster than the next fastest conditions in Exp. 1B (conditions A 'NoPrag-NoPT' and B - 'Prag-NoPT', both at 1604 ms). A regression model on Exp. 1B reaction times including this time condition F - 'Attentional control', and with this condition set as baseline shows that condition is now a significant factor ( $\chi^2(4) = 3.03$ ,  $p < 0.001$  - see Appendix 2), and that it is condition F - 'Attentional control' which is significantly faster than every other condition. This strongly suggests that when participants are selecting the card in the shaded area, they are not doing so after having overcome a -hypothetical- suppression/inhibition strategy for this card.

Most importantly, if participants were indeed using a suppression/inhibition strategy for the card in the shaded area, then the rates of erroneously choosing this card ought to be stable in all of the main conditions, A-E, of the two experiments. The results however show that the rate of selection of this card varies, from near zero in condition A - 'NoPrag-NoPT' to 25% in condition D - 'Prag-PT clash'. This again suggests that participants are not using a shallow strategy of ignoring the shaded card. To the contrary of what such a shallow strategy would suggest, we see that it is the content of the card in the shaded area that predicts the number of erroneous selections, which supports the view that participants are engaging with the card, and reasoning (sometimes correctly, sometimes erroneously) about its availability as a plausible referent via perspective taking. This is exactly what we see in the already well-established literature on referential communication and perspective-taking. Indicatively, in [Heller et al. \(2008\)](#), it is reported that when the object in the privileged position is a competitor (e.g., a duck) to the object in the common ground (e.g., another duck), participants are twice as likely to look to the privileged ground object compared to when it is not a competitor (e.g., a bar of soap). Therefore, both from the research presented here and from a large body of previous

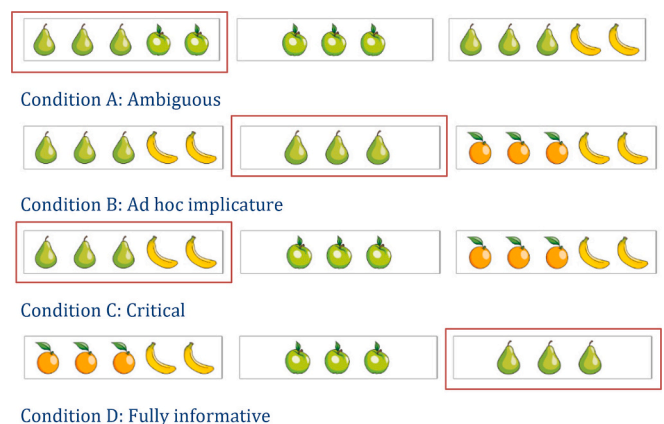
work on the referential communication paradigm we can be confident that participants are not simply ignoring the card in the shaded area.

Of course, at the single-trial level, when perspective-taking suggests that an otherwise preferred card is not a possible referent of the speaker's instructions, the non-selection of this card can be implemented via a number of possible psycholinguistic mechanisms, whether it is by inhibiting that card as a likely referent or by raising the salience of any other card in the common ground etc.; what perspective-taking gives us -in terms of explanatory power for these studies- is the motivation for the participants not choosing the shaded card.

Another concern that might arise is whether the relatively low performance in condition D - 'Prag-PT clash' is not due to the factors we attributed it to but to the fact that the participant who takes the perspective of the speaker has to select a card with two types of items (e.g., with pears and bananas) when the instructions only mention one type of item ('card with pears'). In other words, taken out of context, there is a concern that the instructions are under-informative for the card that ought to be selected. However, this is not likely to be relevant in our experiments because condition A - 'NoPrag-NoPT' was on purpose amended from [Wilson et al. \(2022\)](#) such that the correct selection is a card with two types of items (pears and bananas) when the instructions only mention one type ('pears'). Performance on condition A - 'NoPrag-NoPT' nevertheless was the highest of all conditions, with 98% accuracy in both experiments. Therefore, the selection of a card with two types of items following instructions which ask for one only is not problematic on its own, as long as the instructions suffice for the unique identification of the card. Besides giving this explanation with reference to the baseline condition A - 'NoPrag-NoPT', we also ran an off-line judgement task to investigate if asking for a card with two types of items via instructions that only refer to one type of item is in any way problematic.

## 4. Experiment 2: acceptability judgement task

The aim of this study was to explore how participants rate an instruction that does not fully describe the content of a card but is still informative when it comes to uniquely identifying the card. In the task, participants saw a display of three cards, drawn from those used in Experiments 1A and 1B, and were asked to rate an instruction to pick up a specific card. Displays presented in [Fig. 4](#). are accompanied by the instruction 'pick up the card with oranges'. Participants were told to imagine a game where one person is the 'instructor', and the other is the 'matcher' and then they were asked to rate how good the instructions were for the matcher to select the target card that was highlighted in a red frame. The matcher could not see the frame, which was there to



**Fig. 4.** A set with one display for each condition in which participants were asked to rate how good the instruction "Pick the card with pears" was for a matcher to select the card highlighted with the red border. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

show to the participants the card that the instructor wanted the matcher to pick.

Participants were asked to use a Likert scale appearing at the bottom of the screen with 4-points on the scale ('bad' – 'kind of bad' – 'kind of good' – 'good') following [Jasbi, Waldon, & Degen \(2019\)](#) about the most informative response scales in implicature rating tasks. Of main interest was how participants would rate the critical condition where the instructions 'pick the card with oranges' suffice for unique identification of the target card, but do not fully describe the card (because the card has oranges and bananas). A condition where the instructions were ambiguous ('pick the card with oranges' in the context of two cards, one with oranges and bananas and one with oranges and pears) was included. Here the instructions did not assist the matcher in selecting the target and the condition functions as a low baseline that ought to be rated as 'bad'. Another condition where the instructions were fully informative ('pick the card with oranges' in the context of just one card which had oranges and nothing else) was included as a high baseline that ought to be rated as 'good'. A condition where the selection of the target card required an ad hoc quantity implicature was also included. No conditions included cards on privileged ground, as this was not relevant for this experiment.

#### 4.1. Method

##### 4.1.1. Participants & procedure

The experiment was created on Gorilla TM Experiment Builder ([Anwyl-Irvine et al., 2020](#)). Twenty-five native speakers of English were recruited online (17 Female, age range 18-64, mean age  $37 \pm 15$ ) via [Prolific Academic Ltd \(2014\)](#) and completed this study from a location of their own choice. Once participants provided information about themselves and consent to take part, they were presented with the instructions for the main task as above in a procedure similar to the one used in Experiments 1A and 1B, with the difference that they saw three displays at a time and they were asked to rate how good the instructions were for selecting the card highlighted inside a red frame. None of the participants who took part in Experiments 1A and 1B took part in Experiment 2 and none of those who took part had a history of neurological or cognitive disorders according to self-report. Ethics approval was given by the Ethics Committee of the Faculty of Modern and Medieval Languages of the University of Cambridge.

##### 4.1.2. Analyses

Since the acceptability rating was on a 4-point Likert scale, it was analysed using ordinal rather than linear regression models. We fitted a cumulative link mixed-effects model to our data, using the `clmm()` function of the ordinal package ([Christensen, 2019](#)). The model included random intercepts for item and participant. Parameters could not be uniquely determined with the fully informative condition included, so we removed this condition from further analysis.

#### 4.2. Results and discussion

The results are presented in [Table 5](#). above. [Fig. 5](#). above visualises the distribution of Likert responses for each condition. An analysis of variance based on mixed ordinal logistic regression indicated a statistically significant effect of Condition. Post-hoc comparisons using emmeans package are shown in [Table 6](#).

As expected, the high-baseline condition, Fully informative, was rated consistently positively, with 100% of responses selecting 'good'. Also as expected, the low-baseline ambiguous condition was rated overwhelmingly negatively, with 86.1% of responses either 'bad' (57.8%) or 'kind of bad' (28.3%). Importantly, the Critical condition was rated overwhelmingly positively, with 98% of responses rating it either 'good' (83.6%) or 'kind of good' (14.4%). These ratings help address the concern that in Experiments 1A and 1B participants avoided choosing the correct card in condition D - 'Prag-PT clash' because the instructions

**Table 5**

Percentages of each type of Likert-scale point per Condition in Experiment 2.

| Condition   | Response     | Percentage |
|---|--------------|------------|
| Condition A - 'NoPrag-NoPT' - Ambiguous                 | Bad          | 57.8       |
|   | Kind of bad  | 28.3       |
|   | Kind of good | 12.5       |
| Condition B - 'Prag-NoPT' – Ad hoc quantity implicature | Good         | 1.4        |
|   | Bad          | 37         |
|   | Kind of bad  | 27.4       |
|   | Kind of good | 27.4       |
|   | Good         | 8.2        |
| Condition C - 'NoPrag-PT' – Critical                    | Bad          | 1          |
|   | Kind of bad  | 1          |
|   | Kind of good | 14.4       |
|   | Good         | 83.6       |
|   | Bad          | 0          |
| Condition D - 'Prag-PT clash' – Fully informative       | Kind of bad  | 0          |
|   | Kind of good | 0          |
|   | Good         | 100        |

mentioned only one type of item whereas the card displayed two types of items. We see here that if the instructions with one type of item suffice for unique identification of the card in context, as they did in condition D - 'Prag-PT clash', participants rate such instruction overwhelmingly positively. Importantly, the analyses show that the Critical condition was rated as significantly better than the low-baseline, the ambiguous condition. Finally, the ad hoc implicature condition was rated somewhat ambivalently, with 2/3 of the responses negative and 1/3 positive. We suggest that this mixed rating reflects the participants preference for the speaker to provide all the information required for unique identification explicitly rather than implicitly via implicature.

#### 5. General discussion

In two experiments we explored something that has not been the direct focus of research to date: the role of the listener's perspective in the derivation of implicature. In line with previous work ([Bergen & Grodner, 2012](#); [Breheny et al., 2013](#); [Goodman & Stuhlmüller, 2013](#)) our results show that listeners often adopt the speaker's perspective and they predominantly do not derive implicatures when these are not licensed by the speaker's perspective (accuracy of 75% or above in Exp 1A and 1B for condition D - 'Prag-PT clash' and over 90% in condition E – 'Prag/PT ambiguity'). However, the results also show that the listener's own perspective has an influence too, especially as seen in condition D - 'Prag-PT clash' where we see 22-25% of implicatures that are not licensed from the speaker's perspective. These findings also hint at additional strategies and cues that listeners might use to revise an interpretation that was originally derived exclusively from their own perspective (cf. the difference between condition D - 'Prag-PT clash' and E). Moreover, in Experiment 2 we ruled out a potential confound in the interpretation of the results on condition D - 'Prag-PT clash' that concerned that the instructions, while informative, did not exhaustively describe what was shown on the target card this condition.

If we were to take the traditional theoretical-pragmatic Gricean view of implicature as a psycholinguistic model, we would conclude that our findings are supportive of the view that listeners largely adopt the speaker's perspective (as noted, accuracy of 75% or above in conditions D - 'Prag/PT clash' and E - 'Prag/PT ambiguity'). However, since this view would predict that no implicatures will be derived without the speaker's perspective being considered, the 22-25% of implicatures in Exp 1A and 1B from the listener's perspective in Condition D - 'Prag-PT clash' presents an important challenge.

Here, we think that the models which postulate a range of pragmatic interpretative strategies that listeners can draw from are at an advantage. These models include a broad range of proposals on pragmatic processing which argue that adult neurotypical speakers have a range of pragmatic processing strategies in their disposal, from egocentric to fully

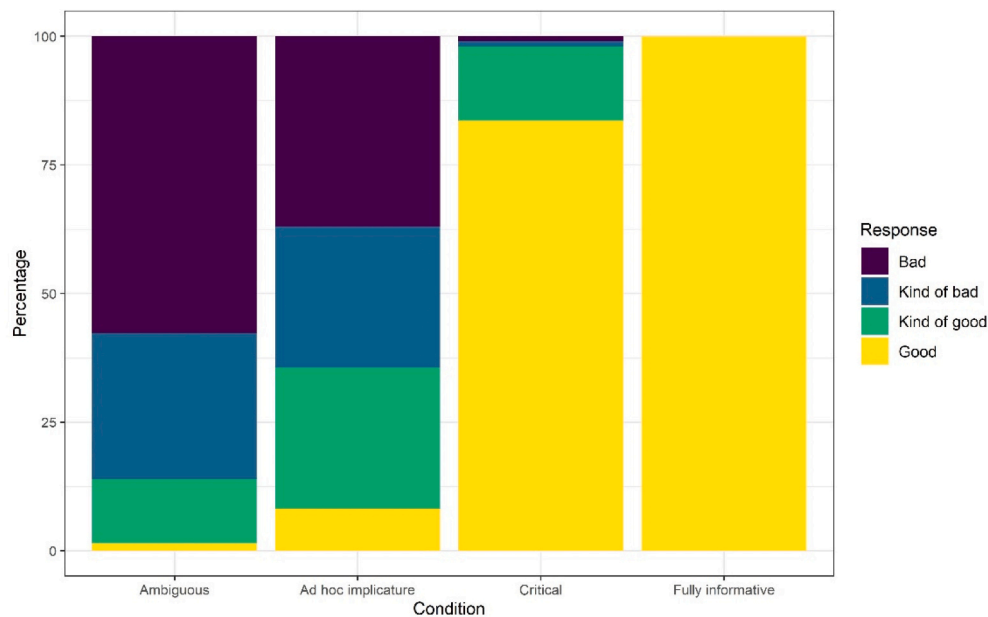


Fig. 5. Distribution of Likert responses per Condition in Experiment 2.

**Table 6**  
Post-hoc analysis output for the model fitting proportions of each type of response in experiment 2.

| Post hoc comparisons – Conditions | $\beta$ | SE   | z-ratio | p value |
|-----------------------------------|---------|------|---------|---------|
| Critical- Ad hoc implicature      | 8.89    | 1.33 | 6.65    | <0.0001 |
| Critical – Ambiguous              | 2.43    | 0.64 | 3.77    | 0.0005  |
| Ad hoc implicature - Ambiguous    | 11.32   | 1.35 | 8.38    | <0.0001 |

Gricean, including a meta-cognitive component which monitors what counts as a successful interpretation in a given task. According to some proposals, reasoning about the speaker's perspective is not always required in pragmatic inferences (see Breheny, 2006; Jary, 2013; Katsos & Andrés-Roqueta, 2021; Kissine, 2016; Moore, 2018; Sperber, 1994; Sperber et al., 2010). For example, Kissine (2016) suggests that pragmatic processes (like implicature derivation) are distinct from pragmatic strategies, which may be more or less egocentric, taking into account the speaker's knowledge to a lesser or greater degree. Depending on the context, hearers may arrive at an interpretation purely egocentrically (based on what is relevant for them in the context), allocentrically (which depends only on first-order Theory of Mind and allows interpretations which are at odds with the speaker's perspective to be ruled out, as is the case with our experiments), or using sophisticated "Gricean" reasoning about the speaker's epistemic state (using second-order Theory of Mind such that the hearer's interpretation of the utterance is embedded in reasoning about the speaker's intentions). Crucially, some sort of pragmatic inferences may be possible under all these strategies in different circumstances, with adults switching between strategies as required (see especially Sperber, 1994 for the first proposal in this direction). To complement this approach, Katsos and Andrés-Roqueta (2021) draw upon work by Clark and Brennan (1991) on grounding and by Garrod and Pickering (2004, Pickering & Garrod 2006) on alignment in conversation to provide criteria for characterizing conversational situations as signals for the listener to abandon ego-centric interpretative strategies and engage in allocentric or fully-Gricean ones.

In addition to more than one interpretative strategy being available, the results call for an explanation as to why accuracy on condition E - 'Prag-PT ambiguity' was higher than on condition D - 'Prag-PT clash'. Here, we hypothesize that the fact that in condition E - 'Prag-PT ambiguity' the strategy which relies exclusively on the listener's perspective

does not lead to an unambiguous selection of a card is a cue for the listener to consider alternative strategies. This could be seen as the outcome of metacognitive processes (Flavell, 1979), with the failure to achieve the desirable goal of unambiguous selection of a card acting as a trigger for engaging metacognitive monitoring in order to adopt a new strategy.

5.1. Relation to pragmatic theory

The studies presented here were inspired by the widely held assumption in pragmatic theory that the speaker's perspective is adopted in the derivation of implicature and asked whether the listener's perspective is important too. While a processing model inspired by Grice's linguistic-philosophical account was not compatible with the evidence that the listener's perspective is important, we saw that processing models that make assumptions about a range of interpretative strategies that speakers have in their disposal, about situations and motivations that favour using one or the other strategy (see Breheny, 2006; Jary, 2013; Katsos & Andrés-Roqueta, 2021; Kissine, 2016; Moore, 2018; Sperber, 1994; Sperber et al., 2010). Nevertheless, all these approaches are inspired by the Gricean approach to inferential communication and the critical role of pragmatic maxims in social interactions. A parallel question concerns the nature of the mechanism that derives pragmatic inferences and whether it is interpersonal in nature or whether it involves a grammatical operator akin to a silent 'only' which triggers the derivation, and then negation, of more informative alternatives (see Geurts, 2010 for a review and critique of such accounts; and Chierchia, 2004; Chierchia et al., 2012 for a defense; among many others). The research presented here does not speak directly to this debate. First, on the grammatical view to implicature, it is not claimed that contextual factors such as the speaker's or the listener's perspective do not play a role, only that the core aspect of the process of deriving an implicature is grammatical in nature. Second, the grammatical proposal capitalises on implicatures from context-independent scales, such as those by logical connectives, modals, quantifiers etc. especially when they embedded under other logical operators. While this account can be extended to ad hoc quantity implicatures and sentences such as the ones used in these experiments, it is not necessary that this is desirable for proponents of the grammatical approach. Overall, we conclude that the debate on the pragmatic vs. grammatical nature of implicature derivation is orthogonal to the

research presented here.

### 5.2. Congruence with previous research on implicature and perspective-taking

Here we return to the issue that in previous research on implicature, with the exception of some indicative findings in [Wilson et al., 2022](#), and ongoing research by [Jarvinen et al., 2022](#), there has been no evidence that listeners are affected on their own perspective. Recall that in the [Breheny et al., 2013](#) study, participants eye-gaze was tracked in conditions where they heard utterances which could be interpreted with an implicature, such as “The woman put a spoon in the box” leading to the implicature “The woman put a spoon in the box and nothing else”. In the critical condition the listener could see that there was a box with a spoon and a box with a spoon and something else. The speaker, however, had not seen the events to their full conclusion because of a screen which blocked her view and therefore she only knew that there were two boxes, each with just a spoon. Listeners did not interpret the instructions with an implicature, looking at either box rather than at the box that was, from their own perspective, compatible with an implicature interpretation. Importantly, however, there is a temporal aspect in this task, with events unfolding over time, and a very salient marker of the epistemic states of the speaker-listener diverging in some occasions only, i.e., the lowering of the screen which prevented the speaker from seeing the final actions that took place. Given proposals from the referential communication paradigm that strong cues for perspective-taking can influence listeners to switch out of an egocentric interpretative strategy as well as to increase their motivation to engage with perspective-taking ([Bezuidenhout, 2013](#); [Brown-Schmidt & Hanna, 2011](#); [Sperber et al., 2010](#)), it is possible that the salient lowering of the screen which occurred in only some trials acted as a cue to the listener to engage with more high-order interpretative strategies in those trials. There was no such strong cue in our experimental paradigm, however, where in every display that the participants saw there was reliably one card that the speakers could not see. The fact that there was no salient cue to signal that some trials require a different interpretative strategy than others might explain why we were able to discern an effect from the listener's perspective in our experiments.

Moreover, we should note that even if some listeners, on some occasions, were affected by what was in their own privileged views in the [Breheny et al. \(2013\)](#) study, the results of eye-tracking patterns were presented at group level, which might conceal the effect.

Turning to the other study that did not reveal effects of the listener's perspective, [Kampa and Papafragou \(2020\)](#), recall that in this study participants heard an utterance and they were asked to decide which display it referred to, the one where the speaker had full knowledge or the one where she had partial knowledge. Adults were at ceiling with matching the speaker's epistemic state with the utterance they made, despite there being conflicting information from the listener's perspective. However, the results are open to interpretation as to what exactly they measure. In this task it is not strictly necessary to integrate the speaker's perspective into the process of deriving or not an implicature. Instead, participants may be deriving an implicature from the utterance anyway, e.g., they may be interpreting ‘I see a spoon’ as ‘I see a spoon and nothing else’ and then deciding which of the two versions of the speaker, the one who can see a spoon or the one who can see a spoon and a bowl, could have said this. This is related but distinct from incorporating the speaker's epistemic state into the process that leads to a derivation or no derivation of an implicature.

To conclude, we propose that there are methodological as well as conceptual reasons why previous research did not reveal evidence for the role of the listener's perspective in implicature derivation. Together with our studies, the emerging body of work suggests that while listeners routinely take the speaker's perspective into account, this is not unexceptionally the case and that the listener's perspective does play a role too, especially when it leads to interpretations that contrast with the

interpretations reached from the speaker's perspective.

### 5.3. The interaction of the speaker and listener's perspective

A remaining question concerns how the perspective of the listener and the speaker interact in utterance interpretation. In section 3.1 above we outlined models of pragmatic processing that argue that listeners have access to interpretative strategies with or without the speaker's perspective. One possibility is that listeners engage with one or the other strategy, and possibly the less cognitively demanding strategy as a default (see [Kissine, 2016](#); [Sperber et al., 2010](#)). If indeed participants employ an ego-centric strategy, one where they consider that the speaker knows all that they themselves know, participants may occasionally switch to a strategy that considers the speaker's perspective, when there is a metacognitive cue that the interpretation from the current strategy is deficient for the goal of the activity (as could be the case in condition D - ‘Prag-PT clash’ in our Experiments 1A or 1B) or when there is a strong cue or motivation for perspective taking (as in the critical trials in [Breheny et al., 2013](#)).

Another possibility, however, is to take a constraint-based approach of implicature, where the speaker's knowledge does have an influence on the probability of an implicature being derived, but it is only one of many other factors, such as the question under discussion, world knowledge, the cost of deriving the stronger alternative or the properties of the utterance itself ([Degen & Tanenhaus, 2015, 2019](#)). Indeed, an additional constraint, as our studies show, would be the listener's own perspective.

Both of the views on implicature and perspective-taking above have parallels in the literature on referential communication, where adopting the speaker's perspective can be seen from a constraint-based approach (e.g., [Brown-Schmidt & Hanna, 2011](#)) or as a matter of selecting one of the two, speaker- or listener-oriented interpretative strategies that are available and occasionally revising it (e.g., [Keysar et al., 2000](#)).

However, there is other work in the referential communication paradigm that proposes that both strategies are active during a single process of interpretation. [Heller et al. \(2016\)](#) and [Heller \(2020\)](#) argue that the choice of a referent requires the incremental and simultaneous integration of the listener's own perspective with the speaker's perspective, each with their own weight, rather than simply being based on one of the two strategies. In this view, a major challenge for listeners would be the integration of the two perspectives when they diverge.

While our findings make the modest contribution that the listener's perspective is important in implicature derivation, we suggest that further research could explore if the speaker's and listener's perspective are two distinct interpretative strategies available to listener's or if a weighted constraints-based view and/or integration view is on the right track. Recall that an important finding was that the listener's perspective has a discernible effect on accuracy rates only when the referent that would be selected from this perspective is different to the referent that would be selected from the speaker's perspective (condition D - ‘Prag-PT clash’). Moreover, the listener's perspective did not seem to affect accuracy when the interpretation from the listener's perspective lead to two possible referents, one of which was the preferred one from the speaker's perspective (condition E - ‘Prag-PT ambiguity’). In Section 2.3 we tentatively attributed the difference in accuracy between these conditions to the engagement of metacognitive process which signal that a different interpretative strategy is needed in condition E - ‘Prag-PT ambiguity’, but not in condition D - ‘Prag-PT clash’.

However, this difference could also be explained as an outcome of the differential way in which the two perspectives must be integrated into a single interpretation. This is especially so when we consider that when the listener's perspective led to two possible referents while the speaker's perspective narrowed these down to one of these two (condition E - ‘Prag-PT ambiguity’), accuracy was at very high levels. In this condition, the integration of the two perspectives is a matter of weighting the cues in favour of each of the two possible referents, with

the referent card that is available both from the speaker's and the listener's perspective being preferred to the referent card that is available only from the listener's perspective. Correspondingly, the lowest performance was evident when each perspective suggested a different referent card (condition D - 'Prag-PT clash') making the integration of the two perspectives a matter of completely rejecting the perspective with the lower weight (the listener's) in order to adopt the perspective with the higher weight (the speaker's). An interesting avenue for future research could be to adapt Heller et al. (2016) computational model of reference with noun phrases to implicature and perspective-taking, and to test the resulting model predictions in ways that can yield information about online processing, such as eye-tracking.

## Data availability

Data available at: <https://osf.io/gnxjb/>

## Acknowledgement

We would like to thank Daphna Heller for insightful discussions on an earlier version of this manuscript as well as Susan A. Graham and the participants of the 'MK40: Common Knowledge, Common Ground, and Context in Communication' conference in London in 2021 for helpful comments. The support of the following research funding bodies is gratefully acknowledged by the authors: Napoleon Katsos was supported by a UK Arts and Humanities Research Council grant (Multilingualism: Empowering Individuals, Transforming Societies, AH/N004671/1), Blanche Gonzales de Linares was funded by an Interdisciplinary PhD scholarship by the UK's Economic and Social Research Council, Ekaterina Ostashchenko was funded by a Wiener Anspach Foundation Post-doctoral Research Fellowship, and Elspeth Wilson was funded by a UK Economic and Social Research Council Post-doctoral Research Fellowship (grant ES/S010203/1).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cognition.2023.105582>

## References

- Allan, K. (2013). Common ground. In J. Verschueren, & J.-O. Ostman (Eds.), *Handbook of pragmatics: 2013 installment* (pp. 1–26). John Benjamins Publishing Company.
- Andrés-Roqueta, C., & Katsos, N. (2020). A distinction between linguistic- and social-pragmatics helps the precise characterisation of pragmatic challenges in children with Autism Spectrum Disorders and Developmental Language Disorder. *Journal of Speech, Language and Hearing Research*, 63(5), 1494–1508. <https://doi.org/10.1044/2020.JSLHR-19-00263>
- Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, 52(1), 388–407. <https://doi.org/10.3758/s13428-019-01237-x>
- Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A., & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *The Quarterly Journal of Experimental Psychology*, 63(6), 1201–1217.
- Barner, D., Hochstein, L. K., Rubenson, M. P., & Bale, A. (2018). Four-year-old children compute scalar implicatures in absence of epistemic reasoning. In K. Syrett, & A. Arunachalam (Eds.), *Vol. 24. Semantics in language acquisition. Trends in language acquisition research* (pp. 325–350). Amsterdam: John Benjamins.
- Barr, D. J. (2008). Pragmatic expectations and linguistic evidence: listeners anticipate but do not integrate common ground. *Cognition*, 109(1), 18–40.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- 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.
- Bezuidenhout, A. (2013). Perspective taking in conversation: A defense of speaker nonegocentricity. *Journal of Pragmatics*, 48(1), 4–16.
- Bott, L., & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language*, 51(3), 437–457.
- Breheny, R. (2006). Communication and folk psychology. *Mind & Language*, 21(1), 74–107.
- Breheny, R., Ferguson, H. J., & Katsos, N. (2013). Taking the epistemic step: Toward a model of on-line access to conversational implicatures. *Cognition*, 126(3), 423–440.
- Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? *Cognition*, 100(3), 434–463.
- Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review*, 16(5), 893–900.
- Brown-Schmidt, S., & Hanna, J. E. (2011). Talking in another person's shoes: Incremental perspective-taking in language processing. *Dialogue & Discourse*, 2(1), 11–33.
- Carston, R. (1998). Informativeness, relevance and scalar implicature. In R. Carston, & S. Uchida (Eds.), *Relevance theory: Applications and implications* (pp. 179–236). Amsterdam: John Benjamins.
- Chierchia, G. (2004). Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and beyond* (pp. 39–103). Oxford University Press.
- Chierchia, G., Fox, D., & Spector, B. (2012). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In P. Portner, C. Maienborn, & K. von Stechow (Eds.), *Vol. 3. Semantics: An international handbook of natural language meaning* (pp. 2297–2331). Berlin: Mouton de Gruyter.
- Christensen, R. H. B. (2019). Ordinal - regression models for ordinal data. R Package Version 2019.4-25. Available at: <http://www.cran.r-project.org/package=ordinal/>.
- Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Behrend (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: American Psychological Association.
- Degen, J., & Tanenhaus, M. K. (2015). Processing scalar implicature: A constraint-based approach. *Cognitive Science*, 39(4), 667–710.
- Degen, J., & Tanenhaus, M. K. (2019). Constraint-based pragmatic processing. In C. Cummins, & N. Katsos (Eds.), *Handbook of Experimental Semantics & Pragmatics* (pp. 21–38). Oxford University Press.
- Del Sette, P., Bambini, V., Bischetti, L., & Lecce, S. (2020). Longitudinal associations between theory of mind and metaphor understanding during middle childhood. *Cognitive Development*, 56. <https://doi.org/10.1016/j.cogdev.2020.100958>. Article 100958.
- Deliens, G., Antoniou, K., Clin, E., & Kissine, M. (2017). Perspective-taking and frugal strategies: Evidence from sarcasm detection. *Journal of Pragmatics*, 119, 33–45.
- Deliens, G., Antoniou, K., Clin, E., Ostashchenko, E., & Kissine, M. (2018). Context, facial expression and prosody in irony processing. *Journal of Memory and Language*, 99, 35–48.
- Dieuleveut, A., Chemla, E., & Spector, B. (2019). Distinctions between primary and secondary scalar implicatures. *Journal of Memory and Language*, 106, 150–171. <https://doi.org/10.1016/j.jml.2019.02.008>
- Epley, N., Morewedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, 40(6), 760–768.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906–911.
- Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, 336(6084), 998.
- Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Science*, 8, 8–11.
- Geurts, B. (2010). *Quantity implicatures*. Cambridge University Press.
- Geurts, B., & Rubio-Fernández. (2015). Pragmatics and processing. *Ratio*, 28(4). <https://doi.org/10.1111/rati.12113>
- Goodman, N. D., & Stuhlmüller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science*, 5(1), 173–184.
- Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. Morgan (Eds.), *Vol. 3. Syntax and semantics* (pp. 41–58). New York: Academic Press.
- Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, 49(1), 43–61.
- Happé, F. G. (1993). Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition*, 48(2), 101–119.
- Hawkins, R. D., Gweon, H., & Goodman, N. D. (2021). The division of labor in communication: Speakers help listeners account for asymmetries in visual perspective. *Cognitive Science*, 45(3), Article e12926. <https://doi.org/10.1111/cogs.12926>
- Heller, D. (2020). The production and comprehension of referring expressions: Definite description. *Language and Linguistics Compass*, 14(5), Article e12370. <https://doi.org/10.1111/lnc3.12370>
- Heller, D., Grodner, D., & Tanenhaus, M. K. (2008). The role of perspective in identifying domains of reference. *Cognition*, 108(3), 831–836.
- Heller, D., Parisien, C., & Stevenson, S. (2016). Perspective-taking behavior as the probabilistic weighing of multiple domains. *Cognition*, 149, 104–120.
- 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.
- Hochstein, L., Bale, A., Fox, D., & Barner, D. (2016). Ignorance and inference: do problems with gricean epistemic reasoning explain children's difficulty with scalar implicature? *Journal of Semantics*, 33(1), 107–135.

- Horn, L. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schriffen (Ed.), *Meaning, form, and use in context: Linguistic applications* (pp. 11–42). Washington, DC: Georgetown University Press.
- Huang, Y. T., & Snedeker, J. (2009). Online interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. *Cognitive Psychology*, 58(3), 376–415.
- Jarvinen, H., O'Shea, K., & Barr, D. (2022). Do listeners suspend pragmatic inferences when perspectives mismatch? A visual-world eyetracking study on scalar adjectives. Poster presented in XPrag 2022 conference, Pavia, 22-23 September 2022.
- Jary, M. (2013). Two types of implicature: Material and behavioural. *Mind & Language*, 28(5), 638–660.
- Jasbi, M., Waldon, B., & Degen, J. (2019). Linking Hypothesis and Number of Response Options Modulate Inferred Scalar Implicature Rate. *Frontiers in Psychology*, 10, 189. <https://doi.org/10.3389/fpsyg.2019.00189>
- Kampa, A., & Papafragou, A. (2020). Four-year-olds incorporate speaker knowledge into pragmatic inferences. *Developmental Science*, 23(3), Article e12920.
- Katsos, N. (2008). The semantics/pragmatics interface from an experimental perspective: The case of scalar implicature. *Synthese*, 165, 385–401.
- Katsos, N., & Andrés-Roqueta, R. (2021). Where next for pragmatics and mind-reading? A situation-based view (Response to Kissine). *Language*, 97(3), e184–e197. <https://doi.org/10.1353/lan.2021.0036>
- Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, 11(1), 32–38.
- Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, 89(1), 25–41.
- Kissine, M. (2016). Pragmatics as metacognitive control. *Frontiers in Psychology*, 6, 2057.
- Lecce, S., Ronchi, L., Del Sette, P., Bischetti, L., & Bambini, V. (2019). Interpreting physical and mental metaphors: Is Theory of Mind associated with pragmatics in middle childhood? *Journal of Child Language*, 46, 393–407.
- Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, 69(1), 1–33. <https://doi.org/10.18637/jss.v069.i01>
- Levinson, S. C. (2000). *Presumptive meanings*. Cambridge, Massachusetts: MIT Press.
- Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology*, 46(3), 551–556.
- Moll, H., & Meltzoff, A. N. (2011). Perspective-taking and its foundation in joint attention. In J. Roessler, H. Lerman, & N. Eilan (Eds.), *Perception, causation, and objectivity* (pp. 286–304). Oxford: Oxford University Press.
- Moore, R. (2018). Gricean communication, language development, and animal minds. *Philosophy Compass*, 13(12), Article e12550.
- Papafragou, A., Friedberg, C., & Cohen, M. (2018). The role of speaker knowledge in children's pragmatic inferences. *Child Development*, 89(5), 1642–1656.
- Pickering, M. J., & Garrod, S. (2006). Alignment as the basis for successful communication. *Research on Language and Computation*, 4, 203–228.
- Prolific Academic Ltd. (2014). *Prolific Academic. Prolific*. [www.prolific.co](http://www.prolific.co).
- R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Ryskin, R. A., Brown-Schmidt, S., Tullis, J., & Benjamin, A. S. (2015). Perspective-taking in comprehension, production, and memory: An individual differences approach. *Journal of Experimental Psychology: General*, 144, 898–915.
- Sperber, D. (1994). Understanding verbal understanding. In J. Khalfa (Ed.), *Vol. 1. What is intelligence* (pp. 179–198). Cambridge: Cambridge University Press.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, 25(4), 359–393.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition*. Oxford: Blackwell.
- Spychalska, M., Reimer, L., Schumacher, P. B., & Werning, M. (2021). The cost of the epistemic step: Investigating scalar implicatures in full and partial information contexts. *Frontiers in Psychology*, 12, Article 679491. <https://doi.org/10.3389/fpsyg.2021.679491>
- Stalnaker, R. (1978). Assertion. In P. Cole (Ed.), *Vol. 9. Syntax and semantics: Pragmatics*. Academic Press.
- Tarricone, P. (2011). *The taxonomy of metacognition*. Psychology Press.
- Wardlow, L. (2013). Individual differences in speakers' perspective taking: The roles of executive control and working memory. *Psychonomic Bulletin & Review*, 20(4), 766–772.
- Wilson, E., Lawrence, R., & Katsos, N. (2022). The role of perspective-taking in children's quantity implicatures. *Language Learning and Development*. <https://doi.org/10.1080/15475441.2022.2050236>
- Wilson, E., Lawrence, R., & Katsos, N. (2023). The role of perspective-taking in children's quantity implicatures. *Language Learning and Development*, 19(2), 167–187. <https://doi.org/10.1080/15475441.2022.2050236>