Running Head: IGNORANCE & INFERENCE

Ignorance and inference: do problems with Gricean epistemic reasoning explain children's difficulty with scalar implicature?

Lara Hochstein[1], Alan Bale[2], Danny Fox[3,4], and David Barner[1]

[1]University of California, San Diego

[2]Concordia University

[3]The Hebrew University of Jerusalem

[4]Massachusetts Institute of Technology

Abstract

Unlike adults, children as old as 9 years of age often fail to infer that a sentence like, "Some of the children slept" implies the falsity of its stronger alternative, "All of the children slept" – an inference referred to as a "scalar implicature". Several explanations have been proposed to account for children's failures with scalar implicature, including domain-general processing limitations, pragmatic deficits, or an inability to access the relevant alternatives in a lexical scale (e.g., *all* as an alternative to *some*). Our study focused on the role of Gricean epistemic reasoning in children's failures by testing their ability to compute "ignorance implicatures", which require reasoning about speaker knowledge and informativeness but which differ from scalar implicature with respect to the alternative statements that are involved. We administered two matched tasks to 4- and 5-year-old children: one that assessed their ability to compute ignorance implicatures, and another that assessed their ability to compute scalar implicatures. Five-year-olds successfully computed ignorance implicatures despite failing to compute scalar implicatures, while 4-year-olds failed at both types of inference. These results suggest that 5-year-olds are able to reason about speaker knowledge and informativeness, and thus that it is difficult to explain their deficit with scalar implicature via these factors. We speculate about other possible sources of their difficulties, including processing limits and children's access to the specific scalar alternatives required by scalar implicature.

1. INTRODUCTION

Communication often leads to inferences that go beyond the apparent literal meanings of utterances. For instance, from an utterance like (1), we infer that the speaker believes Mary did some of her homework, but not all of it.

(1) Mary did some of her homework.

According to Gricean accounts of pragmatic reasoning (e.g., Grice, 1975; Horn, 1972; Gazdar, 1979; Geurts, 2009, 2010; Russell, 2011), this type of inference – called a "scalar implicature" – stems from the assumption that the speaker is being cooperative: hearers assume that the speaker is giving the appropriate amount of information (Maxim of Quantity; Grice, 1975) and is only making statements that he believes to be true and for which he has good evidence (Maxim of Quality; Grice, 1975). Accordingly, if the speaker believed that Mary did all of her homework, then he would have uttered the more informative alternative sentence in (2).

(2) Mary did all of her homework.

Since the speaker did not utter this stronger statement, the hearer concludes that either the speaker did not know whether (2) was true or false, or, if the speaker is assumed to be knowledgeable/opinionated, that he believed (2) to be false (scalar implicature).

By other accounts, the stronger inference that the speaker believes (2) to be false is not derived from Gricean principles but instead is grammatical in nature (Chierchia, Fox, & Spector, 2012; Chierchia 2004, 2006; Fox 2007, Fox & Hackl 2006; Landman 1998, among others). For example, according to Chierchia, Fox, and Spector (2012), scalar implicatures are computed using a phonologically null operator with a meaning akin to *only,* and thus do not involve pragmatic reasoning. On this account, the sentence in (1) can be interpreted with a meaning similar to (3).

(3) Mary did only some of her homework.

By almost all accounts it is assumed, following Horn (1972), that words like *some* and *all* are represented as formal alternatives which belong to common substitution classes – e.g., *Horn scales*[1]. According to this idea, computing a scalar inference involves generating relevant alternative statements to an utterance via Horn scales and negating the stronger (or non-weaker) alternatives, as delineated in Steps I through IV in Table 1, below. Accounts differ only with respect to the mechanism by which the alternatives are negated, as described in Step IV.

Table 1. *Steps involved in Scalar Implicature computation*

I. Compute the literal meaning of a sentence S

II. Generate a set of alternatives to S, called S_{alt}.

III. Remove the alternatives in S_{alt} that are entailed by the original utterance S and call the remaining set S'.

IV. Strengthen the meaning of S with the negation of all of the members of S'.

a. By Neo-Gricean accounts: For each member *p* in S',

i) First, infer that the speaker does not believe *p* (i.e., $\neg B(p)$, under the assumption that the speaker would have uttered *p* if he had known it to be true.

ii) Second, given evidence that speaker is knowledgeable/opinionated, infer that the speaker believes *p* to be false (i.e., $\neg B(p) \wedge [B(p) \vee B(\neg p)] \rightarrow B(\neg p)$) – this step is known as the "epistemic step" (Sauerland, 2004).

[1] In this paper, we discuss alternatives in terms of Horn scales, for the sake of simplicity, but our discussion applies to other accounts of alternatives as well, such as the one proposed by Katzir (2007).

b. By Grammatical accounts:

i) Strengthening occurs via a grammatical operator, contained in the literal meaning (Step I), which is motivated by evidence that the speaker is knowledgeable/opinionated (see Fox 2007).

The nature of scalar implicature, debated for decades by linguists, has been increasingly subjected to experimental study, both by examining the on-line processes involved in adult comprehension (e.g., Noveck & Posada, 2003; de Neys & Schaeken, 2007; Huang & Snedeker, 2009a), and by investigating how the ability to compute implicatures emerges in acquisition (e.g., Smith, 1980; Noveck, 2001; Papafragou & Musolino, 2003, etc). Acquisition offers a particularly useful window into understanding the nature of implicature, because it promises the possibility of dissociating its different representational components. To the extent that epistemic reasoning, verbal working memory, and the semantic representation of scalar items change during the early years of language acquisition, the relative contribution of each factor to implicature can be assessed using a developmental approach.

Unlike adults, young children often judge relatively uninformative statements to be acceptable under circumstances that merit the use of a more informative alternative utterance. Studies by Smith (1980), Noveck (2001), Papafragou and Musolino (2003), and others have shown that children as old as 9 are more willing than adults to accept *some* in descriptions where *all* would be more appropriate (see also Barner, Chow, & Yang, 2009; Huang & Snedeker, 2009b; Hurewitz, Papafragou, Gleitman, & Gelman, 2006; Musolino, 2004). For example, when presented with a scene in which three horses appear and then jump over a log, adults overwhelmingly reject the statement, “Some of the horses jumped over the log”, while 5-year-old children almost always accept it (Papafragou & Musolino, 2003). Similarly, children between 3 and 6 years of age are significantly more willing than adults to accept utterances containing *or* in contexts where stronger statements containing *and* are true (Chierchia et al., 2001). For example, children accept the statement, "Every boy chose a skateboard or a bike" to describe a scene in which every boy chose *both* objects (see also Braine & Rumain, 1981; Paris, 1973).

These observed deficits with scalar implicature are important, since they open a unique window into the nature of the semantics / pragmatics interface. If children exhibit a general inability to compute scalar implicatures, then it may be possible to isolate the basic semantic meaning of utterances as distinct from the enriched meaning that is derived once scalar implicatures are computed. Specifically, this can be achieved by asking how the relevant utterances are interpreted early in acquisition. For example, Papafragou and Musolino (2003) argued that because children are unable to compute scalar implicatures for sentences containing scalar terms like *some* and aspectual verbs like *start*, they must not derive exact meanings of numerals like *two* and *three* via implicature, but instead must have exact lexical meanings for numerals. Critically, this reasoning assumes that the difficulty children have with *some*/*all* and *start/finish* implicatures is sufficiently general to also affect implicatures with all other scales, including numerals. However, as noted by subsequent studies, it is conceivable that children are able to make the computations required by scalar implicature in the general case, and that the difficulties they exhibit for cases like *some/all* are specific to these particular items. For example, given that children memorize numerals as members of a class of alternatives prior to learning their meanings (Carey, 2009; Wynn, 1990, 1992), accessing these number words as relevant alternatives may be easier than accessing members of other scales, like *some/all* (for discussion, see Barner & Bachrach, 2010; Barner, Brooks, & Bale, 2011). This raises the possibility that children have the general capacity to compute implicatures but that this ability is restricted by access to particular scales, such that exact numeral meanings are derived pragmatically in acquisition, even though other scalar items are not exhaustified. More generally, to understand the inferences that we can draw from children's non-adult behavior, we must first understand what causes these deficits, and determine whether they are indeed diagnostic of a general divide between semantics and pragmatics.

In the present study, we investigated the nature of children's delay with scalar implicature by focusing on one particular factor: children's ability to compute Gricean epistemic inferences. In the experiments reported below, we test the case study of disjunction and find, like previous studies, that 4- and 5-year-old children fail to compute scalar implicatures (see e.g., Chierchia et al., 2001; Braine & Rumain, 1981; Paris, 1973). However, we also show that 5-year-olds, but not 4-year-olds, *can* compute a related form of inference, called "ignorance implicature". Ignorance implicature, which we describe in detail below, is an interesting comparison case because it requires the ability to reason about a speaker's epistemic state on the basis of utterance informativeness.

Based on our finding that children compute ignorance implicatures earlier than scalar implicatures, and based on data from past studies, we argue that the cause of children's delay with scalar implicature cannot be epistemic in nature. We therefore consider the other two possibilities that have been suggested in the literature – namely, that children's difficulties with scalar implicatures are due to (1) deficits in general processing capacity, or (2) problems accessing the specific alternatives that are required to compute scalar implicature. In parallel, we consider how different theories of implicature might account for our data, including Neo-Gricean (Sauerland, 2004; Franke, 2011) and grammatical approaches (Chierchia, Fox, Spector, 2009;

Fox, 2010). Although we do not argue for one approach over the other on the basis of our data, we point out the assumptions that each theory might make to account for our data.

1.1 Past Accounts of Delayed Scalar Implicature in Language Acquisition

Previous studies have proposed three broad classes of explanation to account for children's difficulty with scalar implicature. The first type of account is that children's difficulties with scalar implicature are pragmatic in nature.[2] Noveck (2001), for instance, proposes that children master the pragmatics of scalar terms like *some* and *all* very late in development and initially treat them logically (i.e., such that *some* is acceptable whenever *all* is true). He offers a number of possible explanations for children's early literal interpretation of utterances, including the possibility that they are less skilled at detecting quantity violations or have an expectation of relevance that is "more easily satisfied than that of adults" (p. 186). In a similar vein, Katsos and Bishop (2011) argue that children's failures stem from "pragmatic tolerance"; although children are no less aware of violations of informativeness than adults, they may be more tolerant of them when forced to either accept or reject statements. In support of this hypothesis, Katsos and Bishop offer evidence that children perform better on 3-alternative tasks, which do not require outright rejection of an utterance (e.g., by allowing children to reward speakers with a small, medium, or large strawberry).

[2] The pragmatic accounts described here might be restated within the grammatical theory. Under such a restatement, pragmatic differences between children and adults (e.g. children's "pragmatic tolerance") would lead to differences in preferred grammatical parses. For example, adults, who are pragmatically "intolerant" would prefer a parse with the phonologically null operator that leads to scalar implicatures, because only this would yield a more informative sentence, in accordance with the maxims (see Fox 2007, 2013, for ways in which the maxims might affect the choice of a syntactic parse).

The second broad class of accounts argues that children's problems with scalar implicature are not pragmatic in nature but instead stem from difficulty processing linguistic knowledge. Specifically, some have proposed that children fail to compute implicatures because the inferences require significant processing resources, which children lack in early development (Reinhart, 2004; Chierchia et al., 2001; Pouscoulous et al., 2007). To derive an implicature, children must not only compute the literal meaning of an utterance, but also generate the relevant alternatives to this utterance and implicitly contrast them with the original statement (i.e., Steps I through III in Table 1). Any one of these steps – alone, or taken together – might exceed children's capacity to store and manipulate information in early development. Consistent with this, a number of studies of adult language processing have found that computing scalar implicature requires more time than computing the literal meaning of a sentence (Noveck & Posada, 2003; Bott & Noveck, 2004; Huang & Snedeker, 2009a; but see Grodner et al., 2010 for conflicting evidence) and may require greater computational resources (de Neys & Schaeken, 2007; Marty, Chemla, & Spector, in press).

Finally, according to a third class of accounts, children's difficulty with implicature is caused by a difficulty accessing Horn scales, which stems specifically from a failure to represent relevant alternatives as members of Horn scales – i.e., Step II in Table 1 (see Barner & Bachrach, 2010; Barner, Brooks, & Bale, 2011; Bale & Barner, 2013; Chierchia et al., 2001; Foppolo et al., 2012). Whereas on the processing account mentioned above children's failures with scalar implicature reflect a domain-general capacity limit that (all else being equal) affects all forms of implicature equally, this "access to alternatives" hypothesis proposes that children have difficulty accessing specific scales either because they lack knowledge of which items constitute Horn scales, or because their knowledge of the relationships between scalar items is weak in early development (e.g., such that activating one scalar item does not lead to the rapid activation of scalar alternatives). In support of this idea, Barner et al. showed that children have difficulty accessing *some* and *all* as scale mates even when processing sentences that do not involve pragmatic inference, but that they have little difficulty with corresponding utterances that involve contextually defined scales – which have readily accessible alternatives – instead of stored Horn scales that must be retrieved from memory. For example, when shown three animals sleeping and asked whether *only some* of the animals are sleeping, 4-year-old children in their study frequently replied, "yes", despite the fact that this was literally false (rather than merely implicated). Although interpreting *only some* to mean *not all* does not require any implicature computation (since *only some* entails *not all*), it does require accessing *all* as a relevant alternative and negating it (see Rooth, 1992, among others). Critically, children had no difficulty with corresponding utterances that involved contextual, "ad hoc" alternatives, rather than Horn scales. When children were shown the same image of three animals sleeping, they overwhelmingly denied that "only the cat and the dog" were sleeping (on 86% of trials), since the third animal was also asleep (for additional evidence, see Stiller, Goodman, & Frank, under review; see also Barner & Bachrach, 2010, for an argument that mutual exclusivity tasks require many of the same steps as implicature; see also Verbuk, 2012). Based on this evidence, Barner et al. concluded that children's failure with *only some* despite success with cases like "only the cat and dog" stems from an inability to access *all* as a scale mate of *some*.

1.2 Present Study: The Development of Ignorance Implicature

In the present study, we sought to directly assess children's ability to compute Gricean epistemic inferences – i.e., to reason about speaker knowledge and informativeness – by testing an inference – i.e., "ignorance implicature", which on most accounts requires epistemic reasoning abilities that either exceed those required by scalar implicature (Chierchia, Fox, Spector, 2009; Fox, 2007), or are almost identical in nature (e.g., Sauerland, 2004).

Ignorance implicatures arise when a speaker makes a statement as in (4), and hearers infer that the speaker is ignorant regarding the truth of the stronger alternative statements in (5) and (6). Ignorance implicature can be defined, more generally, as an inference that a speaker believes neither in the truth nor the falsity of a given proposition (as in 7).

(4) Billy went to the bar or the café.

(5) Billy went to the bar.

(6) Billy went to the café.

(7) The speaker, s, is ignorant about the proposition p if and only if it is not the case that s believes p (i.e., ¬B(P)) nor is it the case that s believes not p (i.e., ¬B(¬P)).

According to standard Neo-Gricean accounts, ignorance implicature and scalar implicature involve the same computations and the same formal alternatives (Steps I through IV in Table 1; Grice, 1989; Sauerland, 2004; Spector, 2003, 2006), and differ only with respect to the strength of the final inference, as described below. Specifically, upon hearing a sentence as in (4), the hearer assumes that the speaker is being cooperative and has therefore provided the appropriate amount of information (given the set of alternatives). Based on this, the hearer assumes that if the speaker knew that Billy went to the bar, he would have uttered the statement in (5), “Billy went to the bar”, whereas if he knew that Billy went to the café, he would have uttered the statement in (6), “Billy went to the café.” Since the speaker did not choose either of these stronger and more informative alternative utterances, the hearer concludes that the speaker either lacked evidence regarding their truth (i.e., Step IV-i) or believed them to be false (i.e., Step IV-ii). However, if the speaker believed that either statement was false, then he would not have made the original statement: believing that (4) is true (i.e., that Billy went to one of the two locations), yet (5) is false (i.e., that Billy did not go to the bar) implies that the speaker believes (6) is true (i.e., that Billy went to the café), and thus that he should have just uttered (6); similarly, believing that (4) is true, yet (6) is false implies that the speaker believes that (5) is true. Thus, since the negation of the individual disjuncts generates a contradiction relative to the original statement, the hearer is unable to conclude that the speaker believes the relevant alternatives e.g., (5) or (6) – to be false (Sauerland, 2004; see also, Spector, 2003). The hearer therefore concludes that only ignorance is possible – i.e., that the speaker lacks evidence regarding each individual disjunct and thus does not know whether or not Billy went to the bar, or whether or not Billy went to the café – all he knows is that Billy went to one of these two locations (for discussion, see Fox, 2007; Sauerland, 2004; Spector, 2003).

In the present study, we tested the nature of children's difficulty with scalar implicature by assessing their ability to compute both scalar implicatures and ignorance implicatures. First, we presented 4- and 5-year-old children with disjunctive statements that licensed ignorance implicatures (e.g., "The bunny took a cup or a plate") and asked them to decide which of two characters most likely uttered this statement: one who saw exactly what the animal took or one who was blindfolded while the animal picked an object (and thus was "ignorant" of the actual state of affairs). To preview our results, on these trials, we found that 5-year-olds but not 4-year-olds were able to infer ignorance based on the use of disjunctive statements. This finding suggests that, by 5 years of age, children have sufficient pragmatic and processing ability to make Gricean pragmatic inferences based on the use of a weaker alternative statement.

Second, the same group of 4- and 5-year-olds in our study also completed a matched scalar implicature task, modeled after Chierchia et al. (2001). On these trials, subjects were presented with disjunctive statements that licensed scalar implicatures (e.g., "Each animal has an apple or a strawberry") in situations where the stronger utterance containing *and* was more informative (i.e., when each animal had both an apple and a strawberry). Then, subjects were asked to decide which of two characters most likely uttered the given statement: a smart puppet or a silly puppet. Here, unlike in the *Ignorance Implicature* task, both puppets could see exactly what was happening, thus making it clear that the speaker must be knowledgeable regarding the truth of stronger statements (thereby ruling out a problem with the epistemic step). In this case, both the 4-year-olds and the 5-year-olds failed to compute scalar implicatures, despite easily succeeding at matched control trials. This finding thus replicates previous studies showing that children 5 years of age and younger have difficulty computing scalar implicatures. In the General Discussion, we argue that these results make it unlikely that children's difficulties stem from a deficit in Gricean epistemic reasoning. Based on this, we discuss other possible accounts of children's difficulty with scalar implicature. In particular, we ask whether the difference that we find in children's performance on the two forms of inference can be explained by a difference in processing costs required by each or, instead, by the specific lexical alternatives that each form of inference involves.

2. EXPERIMENT

2.1 Methods

2.1.1 Participants

We tested 22 monolingual English-speaking 4-year-olds (7 females, *M* = 4;4, range = 4;0-4;11) and 21 monolingual English-speaking 5-year-olds, (11 females, *M* = 5;4, range = 5;0-5;11), recruited either by phone or through daycares in the greater San Diego area. Two additional children were excluded; one due to experimenter error and the other due to failure to complete the task.

2.1.2 Procedure & Stimuli

Subjects were seated at a small table directly across from the experimenter and were given two tasks: an *Ignorance Implicature* task and a *Scalar Implicature* task. In each task, subjects watched a scene involving action figures, were presented with a sentence, and were asked to determine which of two characters most likely uttered the given statement. The order of these tasks was counterbalanced across subjects.

2.1.3 Ignorance Implicature Task

In this task, subjects were introduced to two plastic action figures, Farmer Brown and Captain Blue. The experimenter wrapped a blindfold around Captain Blue's eyes, and explained that Captain Blue, "has a blindfold on, so he can't see. He can still hear, but he can't see anything, so he might say things that are funny or not true". Each subject then received 4 warm-up trials followed by 10 total test trials.

On each of the 4 warm-up trials, the experimenter placed a small plastic object on the table (e.g., a toy car or toy cup) while a stuffed animal introduced himself and announced his intention to take an object without naming it (e.g., "It's me, bunny! Look what I'm taking"). On two of the trials, subjects were then presented with a sentence explicitly mentioning a lack of sight (e.g., "The bunny took something. I didn't see what it took") and were asked to determine whether it was Captain Blue or Farmer Brown who uttered this sentence. On the other two trials, subjects were presented with a sentence explicitly mentioning sight (e.g., "I saw the bunny take a plate") and were asked to determine who uttered it. These warm-up trials were designed to confirm that subjects understood the crucial difference between Captain Blue and Farmer Brown and to familiarize subjects with attributing sentences to one or the other.

On each of the 10 test trials, the experimenter placed two small objects on the table while a stuffed animal introduced himself, named both items in front of him, announced his intention to take something, and then took one or both of the objects (e.g. "It's me, bear! Look, a cup and a plate! Look what I'm taking!"). Subjects were then presented with a sentence and asked to determine whether it was Captain Blue or Farmer Brown who uttered this statement (e.g., "Someone said, 'The bear took a plate'. Who do you think said that?"). At the end of each trial, subjects were asked to justify their choices ("Why do you think it was Farmer Brown/Captain Blue?"). These test trials consisted of 5 different types, as shown in Table 2 below, which were presented in one of two counter-balanced orders: Order 1: Ignorance-1, Ignorance-2, True-1, Ignorance-2, False, True-2, True-1, Ignorance-2, False, True-2; Order 2: True-2, False, Ignorance-2, True-1, True-2, False, Ignorance-2, True-1, Ignorance-2, Ignorance-1.

Table 2. *Test trials in the Ignorance Implicature Task*

Condition	**Choices**	**Animal Takes**	**Someone Says…**	**Correct Response**
True-1 (Control)	Cup / Plate	Plate	The bear took a plate.	Seeing doll
True-2 (Control)	Cup / Plate	Cup & plate	The bear took a cup and a plate.	Seeing doll
False (Control)	Cup / Plate	Plate	The bear took a cup.	Blindfolded doll
Ignorance-1 (Critical)	Cup / Plate	Plate	The bear took a cup or a plate.[3]	Blindfolded doll

[3] On *Ignorance-1* trials, the animal sometimes took the item corresponding to the first disjunct (e.g., the bear took a cup and someone said, "The bear took a cup or a plate"), and sometimes

Ignorance-2 (Critical)	Cup / Plate	Cup & plate	The bear took a cup or a plate.	Blindfolded doll

Both of the *True* trials were attributable to the seeing doll (Farmer Brown), as he was the only one who knew exactly what the animal took on each trial (and thus should be preferred over the blindfolded doll, who could only guess). The *False* trials were attributable to the blindfolded doll (Captain Blue), as he was the only one in a position to guess incorrectly. Although the *Ignorance-1* statements (in which the animal took one object) were literally true, we expected that subjects would attribute them to the blindfolded doll if they were able to compute ignorance implicatures[4]. Similarly, although the *Ignorance-2* statements (in which the animal took two objects) were also literally true, we expected subjects to attribute them to the blindfolded doll if they either (1) were able to compute ignorance implicature, or (2) could compute scalar implicature. This trial type therefore allowed us to test whether children might begin to compute scalar implicatures before ignorance implicatures (i.e., in the event that they succeeded at the *Ignorance-2* trials, but failed at the *Ignorance-1* trials). Specifically, subjects could attribute these statements to the blindfolded doll either by reasoning that the speaker must not know which things were taken, or by reasoning that the seeing doll would not use *or* in a scenario in which he knows *and* to be true. Additionally, these *Ignorance-2* trials allowed us to confirm that

took the item corresponding to the second disjunct (e.g., the bear took a plate and someone said, "The bear took a cup or a plate").

[4] One reviewer has suggested that subjects could attribute *Ignorance* trials to the blindfolded doll by simply recognizing that a disjunctive statement is the safest utterance for the blindfolded doll to make if he does not know exactly what the bear took, since this statement guarantees that the blindfolded doll will always be right. However, this statement would be an equally safe choice for the seeing doll. Thus, to conclude that it is the blindfolded doll who is most likely to utter this statement, subjects must understand that this statement expresses ignorance or uncertainty about the truth of each individual disjunct, and that the seeing doll has enough knowledge to utter a stronger alternative statement.

if children attributed *Ignorance-1* statements to the blindfolded puppet it was not simply because they noticed a violation of the Maxims of Relevance or Brevity (Grice, 1975) due to the discrepancy between the number of items taken (i.e., one) and the number of items mentioned (i.e., two) on these trials.

2.1.4 Scalar Implicature Task

This task closely resembled the *Ignorance Implicature* task, but differed crucially in that subjects attributed sentences to smart vs. silly speakers – who both had full knowledge about the scene at hand – rather than to knowledgeable vs. ignorant ones. Thus, in this task the possibility of an ignorance implicature was canceled, as it was made clear that the speaker must be knowledgeable regarding the truth of stronger statements.

At the beginning of the task, subjects were introduced to two stuffed animals: "Smart Puppet", who "always says things that are just right", and "Silly Puppet", who "always says things that are a little weird or silly." Each subject then received four warm-up trials followed by eight total test trials. On each trial, a stuffed animal bear and a stuffed animal cow were placed on the table in front of the subject, facing the smart and silly puppets, and the experimenter placed certain items (including plastic fruit, stickers, and presents) in front of the bear and cow. Subjects were then presented with a sentence describing the scene with the bear and cow and were asked to determine whether it was Smart Puppet or Silly Puppet who uttered this statement.

On two of the four warm-up trials, the experimenter placed a small object (e.g., a strawberry) in front of one of the stuffed animals, and subjects were asked to determine whether it was Smart Puppet or Silly Puppet who uttered the statement, "Each animal has a [strawberry]". On the other two warm-up trials, an object was placed in front of each stuffed animal, and subjects were asked to determine who uttered the statement, "Each animal has a [strawberry]". These warm-up trials were designed to confirm that subjects understood the crucial difference between Smart Puppet and Silly Puppet and to familiarize subjects with attributing sentences to one or the other.

On each of the eight test trials, certain items (including plastic fruit, presents, and stickers) were placed in front of the bear and the cow, and subjects were presented with a sentence and asked to determine whether it was uttered by the Smart Puppet or Silly Puppet (e.g., "Someone said, 'Each animal has an apple or a strawberry.' Who do you think said that?"). At the end of each trial, subjects were asked to justify their choices. These test trials consisted of 4 different types, as shown in Table 3 below, and these trials were presented in one of two counter-balanced orders: Order 1: Scalar, True-1, Underinformative, True-2, Underinformative, True-1, Scalar, True-2; Order 2: True-2, Scalar, True-1, Underinformative, True-2, Underinformative, True-1, Scalar.

Table 3. *Test trials in the Scalar Implicature Task*

Condition	**Object(s) in front of Bear**	**Object(s) in front of Cow**	**Someone says…**	**Correct Response**
True-1 (Control)	Apple	Strawberry	Each animal has an apple or a strawberry.	Smart Puppet
True-2 (Control)	Apple Strawberry	Apple Strawberry	Each animal has an apple and a strawberry.	Smart Puppet
Under informative (Control)	3 Apples	3 Apples	Each animal has 2 apples.	Silly Puppet

Scalar (Critical)	Apple Strawberry	Apple Strawberry	Each animal has an apple or a strawberry.	Silly Puppet

Both of the *True* trials were attributable to the Smart Puppet and the *Underinformative* trials to the *Silly Puppet.* For the underinformative trials we used numbers, since 5-year-olds have no difficulty rejecting underinformative, but true, statements with numbers (see Papafragou & Musolino, 2003 and Barner & Bachrach, 2010, for a discussion of what drives this success). Though literally true, we expected the *scalar-or* sentences to be attributed to the *Silly Puppet* if children were able to compute the implicature that *or* implies *not-both.*

Note that although the critical sentences in both the *Ignorance Implicature* task and the *Scalar Implicature* task were disjunctive statements, the sentences in the *Scalar Implicature* task all contained the quantifier “each” (as in Chierchia et al., 2001). This is because without quantifiers, all disjunctive statements generate an ignorance implicature in addition to the scalar implicature, whereas disjunctive statements containing “each” can also be uttered by a knowledgeable speaker, and thus do not generate ignorance implicatures. Thus, “each” was used to ensure that ignorance implicatures could not drive children’s choice of the Silly Puppet in the *Scalar Implicature* task. The Warm-up trials served to verify that subjects could successfully interpret statements with “each”. Also, as in all previous studies in this literature, we assessed children’s performance on scalar implicature trials by comparing them to matched trials within the same task (the *True-1* trials), which also contained the word “each”, but were literally true. Therefore, our conclusions do not hinge on main effects between task types, but instead on within-task comparisons of critical and control trials.

2.2 Results and Discussion

Prior to analysis, we compared performance on *True* and *False* control trials within each task and found no differences between them for 5-year-olds and only one difference for 4-year-olds.[5] Critically, for all analyses reported below, the same pattern of results was found regardless of whether control trials were combined or analyzed separately. Therefore, to simplify analyses, we combined data for control trials. Figures 1 and 2 display performance on the critical and control trials in both tasks by age.

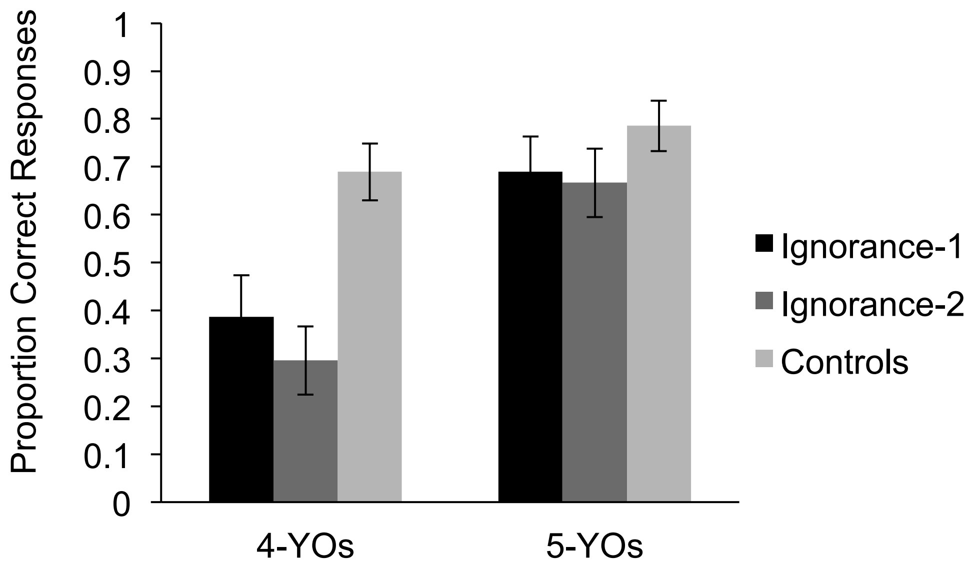

Figure 1. *Proportion of correct responses on critical and control trials in the Ignorance Implicature task. Error bars represent standard error of the mean.*

[5] In the *Scalar Implicature* task, 4 Y.O.s' performance on the *True-2* trials (*M* = .91, *SE* = .05) differed significantly from their performance on the *True-1* trials (*M* = .61, *SE* = .10) and the *Underinformative* trials (*M* = .64, *SE* = .10) (Wilcoxon T = 5, N = 12, *p* = .008, two-tailed, and Wilcoxon T = 0, N = 10, *p* = .005, two-tailed, respectively). However, there was no significant difference between 4 Y.O.s' performance on *True-1* trials and *Underinformative* trials (Wilcoxon T = 38, N = 12, *p* = .95, two-tailed).

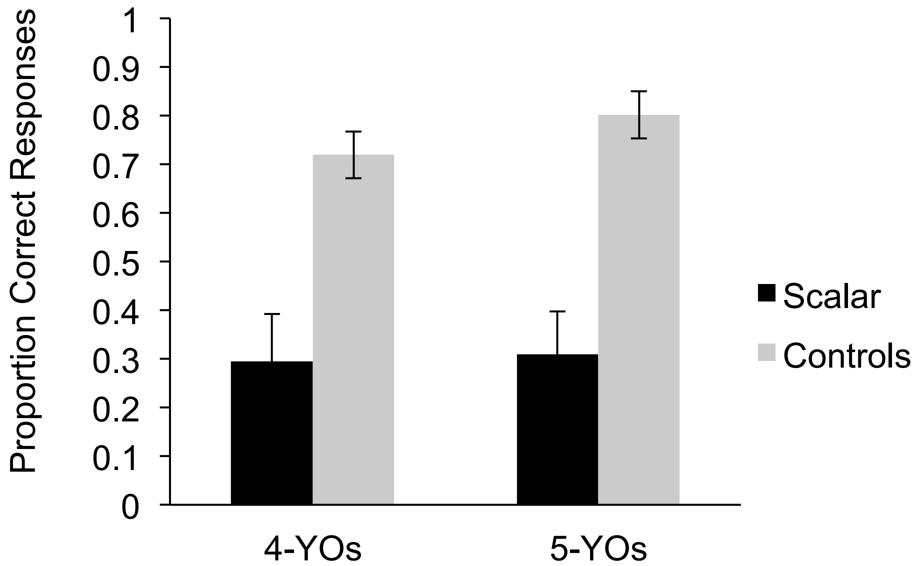

Figure 2. *Proportion of correct responses on critical and control trials in the Scalar Implicature task. Error bars represent standard error of the mean.*

The main question we sought to answer was whether, at any point in development, children exhibit evidence of computing ignorance implicatures before scalar implicatures, or vice versa. To this aim, we initially tested only 5-year-old children, who are known to fail with scalar implicature, and who also are known to have acquired relevant theory of mind capacities. We subsequently tested a separate group of 4-year-olds using the same methods, to determine the age at which ignorance implicatures are first computed for disjunction. Following this logic, we first analyze data for 5-year-olds, then 4-year-olds, and then compare performance across these age groups in a final analysis.

For the 5-year-olds, we conducted separate analyses in which we compared critical *Scalar* trials to either (1) what we called the *Ignorance*-1 critical trials, or (2) what we called the *Ignorance-2* critical trials. First we considered only the *Scalar* and *Ignorance-1 trials* as critical trials, and used a binomial mixed effects regression to predict binary response (correct vs.

incorrect) using Task (Ignorance Implicature vs. Scalar Implicature) and Trial Type (Critical vs. Control) as within-subjects fixed variables and subject as a random factor. This model found an interaction of Task and Trial Type (z = -3.21, *p* = .0013), but no main effects of Task or Trial Type. This interaction was driven by a significant difference between critical (*M* = .31, *SE* = .05) and control trials (*M* = .80, *SE* = .05) for the *Scalar Implicature* task (Wilcoxon T = 11, N = 18, *p* = .0012), but no significant difference between critical *Ignorance-1* (*M* = .69, *SE* = .07) and control trials (*M* = .79, *SE* = .05) for the *Ignorance Implicature* task (Wilcoxon T = 27.5, N = 13, *p* = .22).

Considering *Ignorance-2* trials as critical trials also yielded an interaction effect (z = -3.02, *p* = .0026) but no main effects of Task or Trial Type. Here, as with the *Ignorance-1 trials*, 5-year-olds showed no significant difference between critical trials (*M* = .67, *SE* = .07) and the control trials (*M* = .79, *SE* = .05) for the *Ignorance Implicature* task (Wilcoxon T = 38, N = 15, p = .22). Thus, 5-year-old children preferred to ascribe disjunctive statements to the blindfolded puppet when an animal took either one item (*Ignorance-1* trials) or two items (*Ignorance-2* trials), and in each case did so to the same degree that they responded correctly on control trials. And, critically, 5-year-olds were significantly more likely to correctly attribute disjunctive statements involving ignorance implicature to the blindfolded puppet than they were to correctly attribute disjunctive statements involving scalar implicature to the silly puppet, as shown by the interaction between task and trial type.

To rule out the possibility that the differences found between the scalar trials and the two types of ignorance trials were driven by, for example, particularly strong performance on control trials for the Scalar task, we compared 5-year-olds' performance on the control trials across the two tasks and found no significant difference in performance (Wilcoxon T = 58.5, N = 15, *p* =

.94). This finding also suggests that the two tasks were well matched in difficulty. Relatedly, direct comparisons of critical *Scalar Implicature* trials to *Ignorance-1* and *Ignorance-2* trials each yielded a significant difference (Wilcoxon T = 14, N = 14, p = .016, and Wilcoxon T = 28.5, N = 17, p = .024, respectively).

To explore the nature of these judgments, we asked children to justify their responses following each trial. When asked why they chose the blindfolded doll on *Ignorance-1* and *Ignorance-2* trials, 5-year-olds generally said that it was because Captain Blue (the blindfolded doll) could not see, or because he could still hear what was happening (85.7% on *Ignorance-1* trials and 92.6% on *Ignorance-2* trials). The remaining responses were either unrelated to the question (10.7% on *Ignorance-1* trials and 7.4% on *Ignorance-2* trials – e.g., "because he has white on his hands") – or un-intelligible (3.6% on *Ignorance-1* trials). When 5-year-olds incorrectly attributed *Ignorance-1* and *Ignorance-2* trials to the seeing doll (31% of trials), they typically justified their responses by saying that, "he could see" (83.3% on *Ignorance-1* trials and 91.7% on *Ignorance-2* trials). The remaining justifications were either unrelated (8.3%) or a restatement of the events (8.3% on *Ignorance-1* trials – e.g., "because the bear took the plate"). In contrast, when asked why they chose the smart puppet on the *Scalar* trials (i.e., the non-adult-like response), the 5-year-olds said it was because he was smart (54%), because the statement was right (32%), or because both animals had both objects (14% – e.g., "because they both do"). These justifications suggest that 5-year-olds failed the *Scalar* trials not because they did not understand the distinction between the smart and silly puppets but because they considered sentences like, "Each animal has an apple or a strawberry" to be acceptable descriptions of scenes in which each animal had both objects. The few who correctly attributed *Scalar* trials to the silly puppet justified their responses by saying it was because he was silly (21.4%), because both animals had both objects (14.3%), "because he doesn't know" (14.3%), because the statement was not true (7.1%), or because the speaker used the word "or" (14.3%). The remaining justifications were either unrelated (21.4%) or unintelligible (7.1%).

Our main question was whether children acquire the ability to compute ignorance implicatures before scalar implicatures, and results indicate that they likely do. By the age of 5, children are able to compute ignorance implicatures to the same extent that they respond correctly to true and false control trials. Our next question, then, was exactly how early children are able to compute ignorance implicature. To this aim, we next tested 4-year-olds. As with the 5-year-olds, we conducted two binomial mixed effects regressions to compare performance on the *Scalar Implicature* task first to *Ignorance-1* trials, and then to *Ignorance-2* trials in the *Ignorance Implicature* task. When the *Ignorance-1* trials were treated as critical trials and data compared to results from the Scalar task, we found a main effect of Trial Type ($z = -3.7$, $p = .0002$), indicating that 4-year-olds performed significantly better on control trials than on critical trials in both tasks, but no other significant effects. Similarly, when *Ignorance-2* trials were treated as critical trials, a comparison to the Scalar task found a main effect of Trial Type ($z = -4.61$, $p < .0001$), but no other significant effects. Finally, 4-year-olds showed a significant difference between critical ($M = .30$, $SE = .10$) and control trials ($M = .72$, $ME = .05$) on the *Scalar Implicature* task (Wilcoxon $T = 12.5$, $N = 19$, $p = .0009$), as well as a significant difference between critical *Ignorance-1* trials ($M = .39$, $SE = .09$) and critical *Ignorance-2* trials (M) and control trials ($M = .69$, $SE = .06$), respectively, in the *Ignorance Implicature* task (Wilcoxon $T = 30$, $N = 19$, $p = .009$, and Wilcoxon $T = 25$, $N = 21$, $p = .002$). Thus, unlike 5-year-olds, the 4-year-olds did not perform differently on the ignorance and scalar tasks, and performed significantly better on control trials than on critical trials in each case.

When asked why they chose the seeing doll on *Ignorance-1* and *Ignorance-2* trials (i.e., the non-adult-like response), 4-year-olds said that it was because he could see what was happening (73.9% on *Ignorance-1* trials and 76.9% on *Ignorance-2* trials), "because he knew" (4.3% on *Ignorance-1* trials), or because the blindfolded doll couldn't see (4.3% on *Ignorance-1* trials). The remaining justifications were either unrelated (13.0% on *Ignorance-1* trials and 11.5% on *Ignorance-2* trials) or a restatement of the events (4.3% on *Ignorance-1 trials* and 11.5% on *Ignorance-2* trials). The few who correctly attributed *Ignorance-1* and *Ignorance-2* trials to the blindfolded doll typically justified their responses by saying that "he couldn't see" (80% on *Ignorance-1* trials and 81.8% on *Ignorance-1* trials), that he couldn't see and "thought that [the bunny] took the plate" (6.7% on *Ignorance-1* trials) or that he couldn't see, "so how would he know which?" (6.7% on *Ignorance-1* trials). The remaining justifications were unrelated (6.7% on *Ignorance-1* trials and 18.2% on *Ignorance-2* trials). In contrast, when asked why they incorrectly chose the smart puppet on the *Scalar* trials, the 4-year-olds said it was because he was smart (25%), because the utterance was right (7.1%), because both animals had both objects (42.9% – e.g., "because they both have an apple and a strawberry"), or because he saw it (3.6%). The remaining responses were either unrelated (17.9%) or unintelligible (3.6%). The few who correctly attributed *Scalar* trials to the silly puppet justified their responses by saying it was because he was silly (41.7%), or because both animals had both objects (50.0%). The remaining responses were unintelligible (8.3%).

Our primary question was addressed by our analysis of the 5-year-old children: by the age of 5, children show a clear difference in their ability to compute ignorance implicatures and scalar implicatures. Our follow-up question – when this capacity might first emerge – was probed by testing 4-year-olds, and we found no difference between the two implicature types.

These results suggest that most children acquire the ability to compute ignorance implicatures sometime between the ages of 4 and 5. In our final analysis, we further probed this developmental trajectory by directly comparing the 4- and 5-year-olds.

Using a binomial mixed effects regression, we predicted binary response (correct vs. incorrect) using Age (4 vs. 5) as a between-subjects variable, Trial Type (Critical vs. Control) as a within-subject variable, and subject as a random factor. For the *Scalar Implicature* task, this model found a main effect of trial type ($z = -4.90$, $p < .0001$), indicating that children performed better on control trials than on critical trials overall, but no other significant effects. For the *Ignorance Implicature* task, considering only the *Ignorance-1* trials as critical trials, this model found a main effect of trial type ($z = -3.91$, $p < .0001$), but no other significant effects. Finally, for the *Ignorance Implicature* task, considering only the *Ignorance-2* trials as critical trials, this model found a main effect of trial type ($z = -4.78$, $p < .0001$) and an interaction effect ($z = 2.15$, $p = .03$). Thus, although 5-year-olds clearly compute ignorance implicatures before scalar implicatures, and although only 4-year-olds showed no difference between implicature types, the difference between these two age groups was relatively modest, suggesting that the ability to compute ignorance implicatures emerges quickly between 4 and 5 years of age.

3. GENERAL DISCUSSION

In our experiment, most 5-year-olds successfully computed ignorance implicatures despite failing to compute scalar implicatures, while 4-year-olds generally failed at both types of implicature. Among 5-year-olds, the difference between ignorance and scalar implicature was large and significant, and was found for two different cases of ignorance implicature. In 4-year-olds, however, there was no significant difference between the two inferences, suggesting that children generally acquire the ability to compute ignorance implicatures between the ages of 4 and 5. The relative success of the 5-year-olds on the ignorance implicature task suggests that they have sufficient pragmatic understanding to go beyond the literal meaning of an utterance by reasoning about the role of a speaker's knowledge in his decision to use an underinformative statement. As we argue below, this finding makes it unlikely that children's difficulties with scalar implicature stem purely from a deficit in Gricean epistemic reasoning.

The difference that we report between ignorance implicature and scalar implicature in 5-year-olds is difficult to explain by appealing to pragmatic factors. First, the data are not consistent with a failure to reason about mental states or a lack of sensitivity to informativeness, since ignorance implicature (as argued by Sauerland, 2004 and others) requires the Gricean ability to reason about speaker beliefs and to infer a lack of knowledge on the basis of underinformative statements. Secondly, the data cannot be easily explained by ideas like pragmatic tolerance (Katsos & Bishop, 2011). Unlike previous studies of implicature discussed by Katsos and Bishop, our *Scalar Implicature* task did not require subjects to be "intolerant" and reject statements – which, by the pragmatic tolerance account, children are reluctant to do. Instead, the task merely required children to decide whether the "smart" or "silly" puppet was more likely to utter a given statement. Similarly, if it were true, as Katsos and Bishop argue, that children fail at pragmatic tasks whenever they are binary and require choosing between two alternatives – whether they be statements, puppets, or two evaluations of an utterance – then we would have expected children to fail not only on the *Scalar Implicature* task but also on the *Ignorance Implicature* task, since both were binary judgment tasks. However, most 5-year-olds in our study had no difficulty attributing the underinformative statements in the *Ignorance Implicature* task to the ignorant puppet, despite the fact that our task required making a binary, metalinguistic choice. The fact that children typically accepted underinformative statements when they involve scalar implicature but not when they involve ignorance implicature is thus difficult to explain on the pragmatic tolerance hypothesis.

In addition to being incompatible with a deficit to Gricean epistemic reasoning ability, the results that we report are also difficult to explain for theories that appeal purely to domain general processing limitations. As noted in the Introduction, scalar implicature and ignorance implicature involve highly similar processing abilities, and on most accounts differ only in one step – i.e., Step IV, in which hearers negate alternatives by assuming that the speaker is knowledgeable (the "epistemic step" on Neo-Gricean views), or by deploying the grammatical operator (on the grammatical view). There are several reasons to doubt that problems processing this single step could plausibly cause several years of delay between children's ability to compute ignorance implicatures and scalar implicatures. First of all, on Neo-Gricean accounts, the main component of the epistemic step (i.e., Step IV-ii) involves inferring whether or not a speaker is knowledgeable about the truth of a stronger alternative utterance. Yet, in our study, the knowledge states of speakers were provided contextually and did not need to be inferred; on the critical trials, both the smart and silly puppets could clearly see that each animal had both items.[6] It is thus hard to see why children would have had any difficulty taking the epistemic step when the evidence required for taking this step was made readily available to them. Furthermore, even if children failed to take the epistemic step despite being provided with the necessary evidence, we still would not predict the results reported in our study. Specifically, we would still expect children to attribute the critical sentences on the *Scalar Implicature* task to the silly puppet on the

[6] Children's awareness of this speaker knowledgeability was confirmed by control trials, in which they ascribed underinformative statements to the silly puppet and informative true statements to the smart puppet.

basis of the weaker inference that precedes the epistemic step (i.e., ¬B(*p*)). That is, when presented with an utterance like, "Each animal has an apple or a strawberry", even if children could not take the epistemic step to make the stronger inference that it is not the case that the speaker believes each animal had both items, we would still expect them to make the weaker inference (Step IV-i – i.e., ignorance implicature) that the speaker did not know that each animal had both items (since this does not require the epistemic step). Yet, as previously mentioned, both the silly puppet and the smart puppet could clearly see that each animal had both items. Thus, if children had computed ignorance implicature, then we would have expected them to select the silly puppet, since a smart, knowledgeable puppet would not speak as though ignorant. However, as already noted, children in our study overwhelmingly attributed these utterances to the smart puppet and not the silly puppet, suggesting that they did not infer ignorance on these trials, and thus that their problem with scalar implicature is not simply a failure to take the epistemic step (i.e., inferring that the speaker is knowledgeable/opinionated and thus negating stronger alternative utterances)[7]. Consequently, it is difficult to see how a processing limit that is specific to taking the epistemic step could explain our data.

Could processing limits explain our data on the grammatical view? On such accounts, the primary difference between ignorance implicature and scalar implicature is the fact that only scalar implicature involves grammatical exhaustification. This raises the possibility that processing limits impair children's ability to deploy the grammatical operator that underlies implicature. However, as mentioned in the Introduction, previous studies show that children have no difficulty processing sentences that include overt grammatical exhaustivity operators (i.e.,

[7] We suspect that children do not compute ignorance implicatures on these trials for the same reason that they do not compute scalar implicatures – namely, because (as we argue later in the *Discussion*) they do not yet have access to the relevant alternative utterances (e.g., "Each animal has an apple and a strawberry.")

"only"), and succeed with such sentences *years* before they compute scalar implicatures (Barner et al, 2011). There is thus strong evidence that grammatically exhaustified sentences do not pose a particular processing problem to young children, especially not when they are 5 or 6. Also, although children fail to exhaustify a disjunctive sentence in isolation to compute scalar implicatures (e.g., "Every boy chose a skateboard or a bike"), they do not have difficulty holding the identical disjunctive sentence in working memory while simultaneously representing a stronger conjunctive statement (e.g., "Every boy chose a skateboard and a bike"), and judging correctly that the latter statement is stronger (Chierchia et al., 2001). Finally, there is no positive empirical evidence that children's problems with implicature stem from processing limits, whereas there is accumulating evidence that processing limits cannot explain their difficulties, and especially not the relatively subtle difference between ignorance implicature and scalar implicature.

If children's failure to compute scalar implicatures cannot be easily attributed to pragmatic factors or general processing deficiencies, what else can account for their difficulty with this particular inference, in light of their ability to compute ignorance implicature? One possibility, noted in the Introduction, is that children's failure with scalar implicature is due to a difficulty accessing relevant scalar alternatives (Barner & Bachrach, 2010; Barner, Brooks, & Bale, 2011; Chierchia et al., 2001; Foppollo et al., 2012). For example, in the case of scales like *or-and* children may fail to compute scalar implicatures because of an inability to access *and* as a scale-mate to *or* (for discussion see Barner et al., 2011). Although our study did not directly test children's access to alternatives, unlike some previous studies, this approach nevertheless makes predictions that are consistent with the pattern of data that we report. Below, we review two possible ways in which difficulties accessing alternatives might lead to children's difficulties with scalar implicature.

The first possibility is that lexical alternatives are generated differently for ignorance implicature and scalar implicature, and that this difference explains the earlier emergence of ignorance implicature in language acquisition. This idea is most consistent with the grammatical approach to scalar implicature (e.g., Chierchia, Fox, Spector, 2012; Fox, 2007) though recent Neo-Gricean accounts, based in Game Theory, are likely compatible with this idea as well (e.g., Franke, 2011). According to this idea, lexical alternatives may be generated via distinct mechanisms for ignorance implicature and scalar implicature. For example, on the grammatical view whereas ignorance implicatures are Gricean in nature and can draw upon any salient contextual alternatives, scalar implicatures rely instead on a grammatical operator which acts solely on alternatives provided by grammatically specified Horn scales. Accordingly, children may fail at scalar implicature because they are unable to access Horn scales and thus unable to deploy the grammatical operator necessary for this computation.[8] However, since on this view Horn scales are not necessary for ignorance implicatures, children may still successfully compute ignorance inferences on the basis of contextually-available alternatives. This account would explain the pattern of results we report and also make strong predictions about other cases of implicature. For example, without some modification or additional assumptions to allow contextually specified alternatives in some cases, this account would also predict that young children should have difficulty with scalar implicatures that rely on *ad hoc* scales (e.g., see

[8] It is also possible, under this view, that children simply have difficulty deploying the grammatical operator despite having access to the relevant Horn scales. This seems unlikely, however, given that children are able to exhaustify statements with *only* by 3 or 4 years of age (e.g., Goro et al., 2005; Barner et al., 2011).

Hirschberg, 1985).[9] While this question remains relatively unexplored, some preliminary evidence suggests that *ad hoc* implicatures may actually be relatively easy for young children (see Barner et al., 2011; Stiller et al, under review).

A second way in which access to alternatives might impact scalar implicature, which is compatible with Neo-Gricean accounts as well as with the grammatical view, is if children have a more general difficulty accessing Horn scale alternatives (whether for ignorance or scalar implicature) but can successfully access these alternatives when they are made salient in context. On this view, children might have access to alternatives that can be derived from the contextually salient disjuncts of the original statement, but not to the stronger statement derived from the conjunction, which is not contextually salient. To understand this, consider the example described in the Introduction, in (4), and rewritten here in (8):

(8) Billy went to the bar or the café.

(9) Billy went to the bar.

(10) Billy went to the café.

(11) Billy went to the bar and the café.

Here, the alternatives required for ignorance implicature (i.e., 9 and 10) are retrievable from the context (i.e., they are derived from materials contained within the original statement). However, if children require access to the conjunctive statement in (11) to compute scalar implicature, then they may fail because this alternative is not made contextually salient by the original utterance. In sum, if children can deploy only those Horn scale alternatives that are contextually salient, and scalar implicatures involve alternatives that are not provided contextually, then the pattern of

[9] One such modification would be to claim that the grammatical operator can operate on contextually specified alternatives where contextual specification is defined in the following way: "Let *S* be a sentence uttered after a set of sentences *A*. *S′* is contextually specified if and only if *S′* is a member of *A*."

findings we report in our study – and those reported previously in the literature – may be explained by a failure on the part of children to access relevant scalar alternatives.

One interesting feature of this second approach is that it assumes that children do not consider disjunctive alternatives as relevant to computing scalar implicatures. However, note that in the case of disjunction under the universal quantifier it is often possible to use the individual disjuncts to compute a felicitous scalar implicature. Consider, for instance, the statement in (12), taken from our *Scalar Implicature* task:

(12) Each animal has an apple or a strawberry

Here, the negation of the two individual disjuncts implies the falsity of the stronger alternative involving *and*, as shown in (13):

(13) Negation of Disjunct 1: ¬Each animal has an apple.

Negation of Disjunct 2: ¬Each animal has a strawberry.

∴ ¬(Each animal has an apple and a strawberry)

Thus, if children had considered the individual disjuncts as relevant alternatives to the statement in (12) and negated these alternatives, they should have been able to generate the scalar implicature that each animal did not have *both* items – even if they could not generate "Each animal has an apple *and* a banana" as an alternative. Interestingly, this is clearly not what children did, since both the 4-year-olds and the 5-year-olds in our study failed to attribute underinformative disjunctive statements in the *Scalar Implicature* task to the silly puppet. The results from our study thus suggest that, when faced with disjunctive statements involving scalar implicature, 4- and 5-year-olds do not generally negate the disjunctive alternatives. This possibility is not entirely unreasonable given that we know adults sometimes exclude – or "prune" – certain alternatives when computing implicatures. For example, as noted in the

Introduction, negating the disjuncts leads to contradiction in utterances without a universal quantifier (as in 8), with the consequence that most theories now assume these alternatives are excluded from the computation of scalar implicature (e.g., Sauerland, 2004; Spector, 2006; Fox, 2007; Katzir, 2007). Also, negating the disjuncts can result in inferences that are too strong in certain situations. For instance, in a scene in which every animal has a banana but only some have oranges, the disjunctive statement, "Each animal has a banana or an orange" is felicitous, yet negating the first disjunct would lead to the incorrect conclusion that not each animal has a banana. This fact could lead listeners to rule out the use of disjuncts as alternatives for scalar implicature, However, these proposals are speculative, and it remains unclear under what circumstances adults actually exclude particular alternatives from computations.

In sum, to account for the results of our study, any theory of implicature must explain why children generate ignorance implicatures for simple disjunctive statements but do not generate either ignorance or scalar implicatures for disjunctive statements embedded under a universal quantifier. We have argued that these findings are not easily explained by appealing to pragmatic or general processing factors and that, instead, they are best explained by an inability on the part of children to access the specific lexical alternatives on a given Horn scale. Finally, we have shown how both grammatical and Neo-Gricean accounts of implicature can account for these findings, given certain auxiliary assumptions.

Before concluding, it is worth noting that, beyond the specific case study of scalar implicature, our study has broader implications regarding the development of Gricean pragmatic reasoning abilities. The results for 5-year-olds suggest that Gricean epistemic reasoning abilities precede the ability to compute scalar implicatures, a conclusion which is consistent with studies showing that some forms of pragmatic reasoning emerge early in development, such as sensitivity to speaker knowledge and intentions (as noted in the Introduction). However, although some pragmatic abilities emerge early, our finding that 4-year-olds fail to compute ignorance implicatures suggests that Gricean reasoning about utterances may not be fully mature by this age. Computing ignorance implicatures requires not only identifying whether another person is fully or only partially knowledgeable, but also understanding that only someone who lacks knowledge should use a less informative description (i.e., by assuming that a knowledgeable speaker would have used a more informative description). Four-year-olds may still struggle to make inferences about mental states on the basis of subtle linguistic cues like utterance informativeness. And, relatedly, they may initially fail to use speaker knowledge to make judgments of informativeness (for related discussion, see Horowitz & Frank, 2012; Beal & Belgrad, 1990). Thus, although the capacity to reason about intentional states may begin to emerge very early in development, the ability to make inferences about knowledge on the basis of utterance informativeness may take much longer. Future studies should explore this question, and should examine a wider range of test cases, beyond the case of disjunction.

4. CONCLUSION

To summarize, our study provides evidence that many children have the ability to compute relatively sophisticated Gricean epistemic inferences by the age of 5, in the form of ignorance implicatures. We have argued that these findings make it unlikely that epistemic reasoning problems explain children's difficulty with scalar implicature. Instead, we considered two other possibilities: processing limits and access to alternatives. We argued that the weight of the evidence, both from our study and from previous experiments, favors of the hypothesis that children fail at scalar implicature due to problems accessing relevant scalar alternatives. Future studies should investigate how children acquire the ability to compute ignorance implicatures, and how these inferences differ from children's earlier pragmatic abilities.

Author's Address:

Lara K. Hochstein

Department of Linguistics, UCSD

9500 Gilman Drive, La Jolla, CA, 92093

larakhochstein@gmail.com

Acknowledgements

Many thanks to Andy Kehler, Jessica Sullivan, Michael Frank, Mahesh Srinivasan, Joshua Hartshorne, Ellie Chestnut, Katie Kimura, Annalice Heinz, Troy Campbell, Laura Taouk, two anonymous reviewers, members of the UCSD Developmental Brown bag, the Language and Development Lab, and AndyLab, and to all the families that participated in this study. Funding for this project was provided through a fellowship by the Jacob K Javits Foundation and a grant from the James S. McDonnell Foundation.

REFERENCES

Bale, Alan & David Barner (2013), Grammatical Alternatives and Pragmatic Development. In Anamaria Falaus (ed.), *Alternatives in Semantics*, '*Studies in Pragmatics, Language and Cognition'*. Palgrave Macmillan. 238-266.

Barner, David, & Asaf Bachrach (2010), Inference and exact numerical representation in early language development. *Cognitive Psychology*, **60**: 40–62.

Barner, David, Neon Brooks, & Alan Bale (2011), Accessing the unsaid: The role of scalar alternatives in children's pragmatic inference. *Cognition,* **188**: 87-96.

Barner, David, Katherine Chow, and Shu-Ju Yang (2009), Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology,* **58**: 195-219.

Beal, Carol & Susan Belgrad (1990), The development of message evaluation skills in young children. *Child Development*, **61**: 705-712.

Bott, Lewis & Ira A. Noveck (2004), Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of memory and language*, **51**: 437-457.

Braine, Martin & Barbara Rumain (1981), Development of comprehension of "or": Evidence for a sequence of competencies. *Journal of Experimental Child Psychology*, **31**: 46-70.

Carey, Susan (2009), *The origin of concepts*. Oxford University Press, Oxford.

Chierchia, Gennaro (2004), Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In Adriana Belletti (ed.), *Structures and beyond*. Oxford University Press. Oxford. 39-103.

Chierchia, Gennaro (2006), Broaden your views: implicatures of domain widening and the "locality" of language, *Linguistic Inquiry*, **37**: 535-590.

Chierchia, Gennaro, Stephen Crain, Maria Teresa Guasti, Andrea Gualmini, & Luisa Meroni (2001), The Acquisition of Disjunction: Evidence for a Grammatical View of Scalar Implicatures. *Proceedings of the 25th Boston University Conference on Language Development*. Cascadilla Press. Somerville. 157-168.

Chierchia, Gennaro, Danny Fox, & Benjamin Spector (2009), Hurford's constraint and the theory of scalar implicatures. *Presuppositions and implicatures*, 60, 47-62.

Chierchia, Gennaro, Danny Fox, & Benjamin Spector (2012), The grammatical view of scalar implicatures and the relationship between syntax and semantics. In Claudia Maienborn, Klaus von Heusinger, & Paul Portner. (eds.), *Semantics: An International Handbook of Natural Language Meaning*; Volume 1. Mouton de Gruyter. New York.

De Neys, Wim & Walter Schaeken. (2007), When people are more logical under cognitive load: Dual task impact on scalar implicature. *Experimental Psychology*, **54**: 128-133.

Foppolo, Francesca, Maria Teresa Guasti, and Gennaro Chierchia. (2012), Scalar implicatures in child language: Give children a chance. *Language Learning and Development*, **8**: 365-394.

Fox, Danny & Martin Hackl (2006), The universal density of measurement. *Linguistics and Philosophy*, **29**: 537-586.

Fox, Danny (2007), Free choice and the theory of scalar implicatures. In Uli Sauerland and Penka Stateva (eds.), *Presupposition and implicature in compositional semantics*. Palgrave Macmillan. Houndmills, Basingstoke.

Franke, Michael (2011), Quantity implicatures, exhaustive interpretation, and rational conversation. *Semantics and Pragmatics*, **4**: 1-82.

Gazdar, Gerald (1979), *Pragmatics: Implicature, presupposition, and logical form*. Academic Press. New York.

Geurts, Bart (2009), Scalar implicature and local pragmatics. *Mind and Language*, 24, 51-79.

Geurts, Bart (2010), *Quantity Implicatures*. Cambridge University Press. Cambridge.

Goro, Takuya, Utako Minai, & Stephen Crain (2005), "Two disjunctions for the price of only one." *Proceedings of the 29th Boston University Conference on Language Development, Somerville.* Cascadilla Press. MA.

Grice, Paul (1975), Logic and conversation. In P. Cole & J. Morgan (eds.), *Syntax and Semantics,* Volume 3. Academic Press. New York.

Grice, Paul (1989), *Studies in the Way of Words*. Harvard University Press. Cambridge.

Grodner, Daniel, Natalie Klein, Kathleen Carbary, & Michael Tanenhaus. (2010), "Some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition*, **116**: 42-55.

Hirschberg, Julia (1985), *A theory of scalar implicature*. Ph.D. thesis. University of Pennsylvania. Pennsylvania, PA.

Horn, Laurence (1972), On the Semantic Properties of Logical Operators in English. Ph.D. thesis. University of California. Los Angeles, CA.

Horowitz, Alexandra, & Michael Frank (2012), Learning from speaker word choice by assuming adjectives are informative. *Proceedings of the 34th Annual Meeting of the Cognitive Science Society*.

Huang, Yi Ting & Jesse Snedeker (2009a), Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics interface. *Cognitive Psychology*, **58**: 376–415.

Huang, Yi Ting & Jesse Snedeker (2009b), Semantic meaning and pragmatic interpretation in five-year olds: Evidence from real time spoken language comprehension. *Developmental Psychology*, **45**: 1723–1739.

Hurewitz, Felicia, Anna Papafragou, Lila Gleitman, & Rochel Gelman (2006), Asymmetries in the acquisition of numbers and quantifiers. *Language Learning and Development*, **2**: 77–96.

Katsos, Napoleon & Dorothy Bishop (2011), Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*, **120**: 67-81.

Katzir, Roni (2007), Structurally-defined alternatives. *Linguistics and Philosophy*, **30**: 669-690.

Landman, Fred (1998), Plurals and Maximalization. In Susan Rothstein (ed.), *Events and Grammar*. Kluwer. Dordrecht. 237–271.

Marty, Paul, Emmanuel Chemla, & Benjamin Spector (in press), Interpreting numerals and scalar items under memory load. *Lingua*.

Musolino, Julien (2004), The semantics and acquisition of number words: Integrating linguistic and developmental perspectives. *Cognition*, **93**: 1–41.

Noveck, Ira & Andres Posada (2003), Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language*, **85**: 203–210.

Noveck, Ira (2001), When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition,* **78**: 165-188.

Papafragou, Anna. & Musolino, Julien (2003), Scalar implicatures: experiments at the semantics-pragmatics interface. *Cognition,* **86**: 253-282.

Paris, Scott (1973), Comprehension of language connectives and propositional logic relationships. *Journal of Experimental Child Psychology*, **16**: 278-291.

Pouscoulous, Nausicaa, Ira Noveck, Guy Politzer, & Anne Bastide (2007), A developmental investigation of processing costs in implicature production. *Language Acquisition*, **14**: 347-375.

Reinhart, Tanya (2004), The processing cost of reference-set computation: Acquisition of stress shift and focus. *Language Acquisition*, **12**: 109-155.

Rooth, Mats (1992), A theory of focus interpretation. *Natural language semantics*, **1**: 75-116.

Russell, Benjamin (2011), Topics in the Computation of Scalar Implicatures. PhD dissertation, Brown University.

Sauerland, Uli (2004), Scalar Implicatures in Complex Sentences. *Linguistics and Philosophy,* **27**: 367-391

Smith, Carol (1980), Quantifiers and question answering in young children. *Journal of Experimental Child Psychology*, **30**: 191–205.

Spector, Benjamin (2003), Scalar Implicatures: Exhaustivity and Gricean Reasoning. *Proceedings of the ESSLLI*, **3**: 277-288.

Spector, Benjamin (2006), Aspects de la pragmatique des opérateurs logiques. Ph.D. thesis, Université Paris 7, Paris.

Stiller, Alex, Noah Goodman, & Michael Frank (under review), Ad-hoc scalar implicatures in adults and children.

Verbuk, Anna (2012), Developmental evidence against the theoretical distinction between Horn and pragmatic scales. *Journal of Pragmatics*, **44**:1680-1700.

Wynn, Karen (1990), Children's understanding of counting. *Cognition*, **36**: 155-193.

Wynn, Karen (1992), Children's acquisition of number words and the counting system. *Cognitive Psychology*, **24**: 220–251.