Scalar implicature in absence of epistemic reasoning? The case of autism spectrum disorder

Lara Hochstein[1], Alan Bale[2] & David Barner[1,3]

[1]Department of Linguistics, University of California, San Diego

[2]Linguistics Program, Concordia University

[3]Department of Psychology, University of California, San Diego

CONTACT:

David Barner

Department of Psychology, UCSD

9500 Gilman Drive, La Jolla CA

92093-0109

Ph: 858-246-0874

barner@ucsd.edu

Abstract

We investigated “scalar implicature” in adolescents and children with ASD to test whether theory of mind deficits associated with autism affect pragmatic inferences in language. We tested scalar implicature computation in adolescents with ASD (12 – 18 years) and asked whether they reason about mental states when computing inferences. Like previous studies, we found the adolescents with ASD computed implicatures to the same degree as neurotypical adults. However, we also found that this ability may not rely on epistemic reasoning. In a test of epistemic reasoning (which probed so-called “ignorance implicature”) we found that adolescents with ASD were able to make epistemic inferences required by Gricean models of scalar implicature when they were explicitly required by the task to do so. However, in a second task, which asked whether subjects spontaneously reason about mental states in the service of scalar implicature when not explicitly asked to, we found that adolescents with ASD did not engage in epistemic reasoning, leading them to compute scalar implicatures in contexts in which they were not justified. Based on these data, we argue that epistemic reasoning may not be a core, constitutive component of scalar implicature.

INTRODUCTION

When we use language, we often make inferences that appear to go beyond the literal meaning of utterances. For instance, consider the sentence in (1):

(1) Some of the students passed the exam.

(2) All of the students passed the exam.

As a listener, we infer from (1) that some, but not all, of the students passed the exam. This inference is commonly referred to as a "scalar implicature" since the use of the scalar term *some* in (1) implies that the stronger statement in (2), which involves *all* – its "scale mate" – is not true. Scales – like the one formed by the quantifiers *some* and *all* – provide sets of alternative lexical items that express differences in the informational strength of the sentences in which they are used (Horn, 1972). Under circumstances in which the utterance in (2) is true, it is *stronger* than the utterance in (1), even though both would be technically true in such a scenario (e.g., if all students passed the exam).

According to standard Gricean and neo-Gricean accounts of communication (e.g., Grice, 1975; Horn, 1972; Gazdar, 1979), scalar implicature involves the assumption that the speaker is being truthful and cooperative and is providing an appropriate amount of information given what she knows to be true. Thus, in (1) we assume that if all of the students had passed the exam, then the speaker would have uttered the sentence in (2). Because the speaker in (1) chose not to utter the stronger and more informative utterance in (2), we infer that she does not believe it to be true. On such accounts, then, scalar implicature is a form of epistemic reasoning: it involves reasoning about the beliefs expressed in a statement and how they relate to beliefs that might have been expressed by alternative linguistic descriptions, but weren't.

Together, the components of this inference can be described, informally, in three steps:

I. Compute the literal meaning of a sentence S

e.g., Some of the students passed the exam

II. Generate a set of relevant scalar alternatives to S

e.g., All of the students passed the exam

III. Strengthen the meaning of S by negating the relevant stronger alternatives.[1]

e.g., Some, but not all, of the students passed the exam

Critically, the third step – i.e., negating stronger alternatives – can be broken down into two components on the Gricean approach, each of which has an epistemic component not present in the previous two steps. First:

IIIa. For a given alternative statement *p,* infer that the speaker does not believe *p*, under the assumption that the speaker would have uttered *p* if he had known it to be true. – i.e., ¬B(*p*).

So, in example (1), we assume that the speaker does not believe the alternative statement in (2) to be true. This first step is sometimes called an "ignorance implicature" since it involves inferring ignorance on the part of the speaker. Next, this lack of belief regarding the stronger alternative is further strengthened as in IIIb.

IIIb. Given evidence that the speaker is knowledgeable, infer that the speaker believes *p* to be false – i.e., B(¬*p*)).

This goes beyond the assumption that the speaker does not know *p* to be true, and involves an additional assumption on the part of the listener, known as the "Epistemic Step", that the speaker is knowledgeable regarding the truth of the stronger statement ("All of the students passed the exam"), and thus believes it to be false (Sauerland, 2004).

---

[1] Technically, not just "stronger" alternatives, but also those that are not weaker than the original sentence.

Recently, the precise role of epistemic reasoning in scalar implicature has been a topic of substantial debate. By some accounts, although it is possible to characterize implicature as a form of Gricean epistemic reasoning, this characterization is not necessary; implicatures may also be represented as grammatical operations over alternative utterances, which, although conditioned on epistemic assumptions, are not themselves epistemic inferences. For example, according to the grammatical view of implicature, a sentence like the one in (1) is strengthened via a phonologically null grammatical operator, whose content is similar to that of the focus element *only* in English (Chierchia et al., 2012; Fox, 2007). By other accounts, scalar implicatures are computed by default (Levinson, 2000), and may be revised or canceled according to pragmatic considerations, like knowledge that the speaker is in fact ignorant regarding the truth of stronger alternative utterances.

While there has been significant debate within linguistics regarding the role of epistemic reasoning in scalar implicature, there is relatively little empirical evidence to address this question. A number of studies have tested the role of epistemic reasoning by investigating children (Noveck, 2001; Papafragou & Musolino, 2003; Smith, 1980, etc.). This work is sometimes premised on the assumption that children's epistemic reasoning abilities – including false belief reasoning – are weak (Wimmer & Perner, 1983; Wellman et al., 2001). Early studies repeatedly found that children struggle to compute scalar implicatures and similar inferences. However, recent evidence suggests that although children struggle with some aspects of epistemic reasoning, they are relatively sophisticated in other areas, making it unclear whether deficits in epistemic reasoning explain problems with implicature. For example, even young infants demonstrate sensitivity to speaker knowledge and intentions in some contexts (Baldwin, 1991; Tomasello, & Barton, 1994; Sabbagh & Baldwin, 2001; Liszkowski et al., 2008) and by the age of 2 can compute inferences like mutual exclusivity that are similar in structure to scalar implicature (e.g., Clark, 1990; Diesendruck & Markson, 2001; Markman, et al., 2003; for discussion, see Barner & Bachrach, 2010; Bale & Barner, 2012). Based on such observations, some recent studies have argued that children's difficulty with scalar implicature may not lie so much with epistemic reasoning as with other factors, such as the problem of identifying which scalar alternatives are relevant to a given utterance – e.g., that *all* is a scale mate with *some*, and thus should be considered as an alternative when computing implicatures (Barner & Bachrach, 2010; Barner, Brooks, & Bale, 2011; Hochstein, Bale, Fox, & Barner, 2015; Chierchia et al., 2001; Foppollo et al., 2012; Skordos & Papafragou, 2016; Stiller, Goodman, & Frank, 2011).

One important lesson offered by the developmental literature is that pragmatic reasoning is non-monolithic. Evidence for a deficit in one area of reasoning – like theory of mind – does not necessarily mean that children have pragmatic deficits across the board. Therefore, we can't assume that if a child fails a false belief task that they will therefore lack the epistemic abilities to compute a scalar implicature, if such reasoning is indeed required. Unfortunately, while this lesson now seems relatively clear in the case of typically developing children, a similar logic has nevertheless been pursued in another population known to experience a range of pragmatic deficits: i.e., individuals with an autism spectrum disorder (or, ASD). ASD refers to a neuro-developmental disorder characterized primarily by deficits in social communication and interaction as well as repetitive behaviors and interests (APA DSM-V, 2013). Although linguistic abilities vary widely across the autism spectrum, pragmatic deficits generally form a core component of an ASD diagnosis. Even high-functioning individuals with autism (i.e., those with normal cognitive abilities and relatively unimpaired core language abilities) often exhibit difficulties with pragmatic aspects of language such as conversational theory of mind, turn-taking, prosody, irony, humor, and metaphor (see Baron-Cohen et al., 1985; Baron-Cohen, 1988; 1995; Minshew et al., 1995; Baron-Cohen, 1988; Eales, 1993; Frith, 2001; Happé, 1993; Ozonoff & Miller, 1996; Shriberg et al., 2001; Tager-Flusberg, 1999; Tager-Flusberg, Paul, & Lord, 2005).

Although individuals with ASD often exhibit difficulties with various forms of pragmatic reasoning, there is no evidence that they lack the specific epistemic reasoning abilities required by scalar implicature. Despite this, two recent studies have investigated scalar implicature in individuals with ASD, under the assumption that subjects should only succeed if implicature does not require epistemic reasoning. First, in a study by Pijnacker et al. (2009), high functioning adults with ASD were asked to evaluate a set of sentences which the authors argued should involve scalar implicature (these stimuli were similar to those used in previous studies by Smith, 1980; and Noveck, 2001). For example, they were asked whether sentences like “Some sparrows are birds”, and “Zebras have black or white stripes” were “true” or “false”, on the assumption that each should be judged false if subjects computed implicatures. Pijnacker et al. found that the subjects with ASD did not differ from neurotypical controls: both groups provided “false” judgments to the same degree for target sentences. In a similar study, Chevallier et al. (2010) presented high-functioning adolescents with ASD and neurotypical controls with *or* statements that described two pictures on a screen and asked subjects to decide whether these utterances were “right” or “wrong”. For instance, subjects saw a picture of a flower and a frog and were asked to judge the sentence, “There is a frog OR a flower”. Here, again, no significant difference in performance was found between the two test groups; the high-functioning adolescents and adults with ASD were just as likely as controls to reject statements involving scalar implicature that were logically true but pragmatically infelicitous. On the basis of these findings, Pijnacker et al. (2009) and Chevallier et al. (2011) both concluded that high-functioning adolescents and adults with ASD are capable of computing scalar implicatures.

Were we to assume that pragmatic deficits are monolithic, we might conclude from these results that scalar implicature does not strictly require epistemic reasoning, since individuals with ASD are known to struggle with theory of mind, *inter alia*. Some people – and thus possibly all people – may be able compute implicatures in absence of reasoning about the knowledge states of their interlocutors. However, this conclusion is premature for two important reasons. First, similar to the case of typically developing children, it may turn out that the epistemic inferences required by scalar implicature are actually well within the capacity of many individuals with ASD, and that they do in fact reason pragmatically when computing implicatures. Second, it is not clear that the studies by Pijnacker et al. and Chevallier et al. actually tested scalar implicature computation, *per se*. In these studies, subjects were asked to make truth value judgments about sentences – i.e., to judge them as "true" or "false" or "right" or "wrong". This method is potentially problematic for assessing scalar implicature computation since under-informative statements are generally analyzed as being literally true despite being infelicitous (for further discussion of the limits of this method, see Katsos & Bishop, 2011; Papafragou & Tantalou, 2004). For this reason, it is difficult to say whether subjects in these experiments – whether those with ASD or those without – construed the tasks to be tests of felicity, or instead as tests of truth and falsity. Also, relatedly, the tasks used in both studies did not strictly require reasoning about the beliefs of other individuals in context. Instead, subjects were presented with sentences in absence of conversational context either as text on a screen or through headphones in random sequence. In each case, it is unclear whether any reasoning about the intentional states of "cooperative, truthful, speakers" was involved.

In the present study, we investigated the role of epistemic reasoning in scalar implicature by testing two main questions. First, we asked whether individuals with ASD are able to compute the specific epistemic inferences that are thought, on neo-Gricean accounts, to underlie scalar implicature – namely ignorance implicatures. Second, we asked whether these same individuals can compute scalar implicatures, and if so, whether these inferences in fact draw on the epistemic states of interlocutors, or are instead computed via a non-pragmatic route. To answer these questions, we tested subjects using a felicity judgment task to assess scalar implicature, as well as two additional tasks which probed reasoning about epistemic states in the service of implicature.

The first of these tasks tested ignorance implicature. As already noted, ignorance implicature is a form of conversational inference that involves reasoning about epistemic states, and, critical to the present study, is thought by neo-Gricean theories to be a necessary component of scalar implicature (for discussion, see Fox, 2007; Sauerland, 2004; Spector, 2003). Consider, for example, the utterance in (3):

(3) Billy ate pizza or pasta for lunch.

(4) Billy ate pizza.

(5) Billy ate pasta.

From the utterance in (3), the listener should infer that the speaker does not know which of the two foods Billy ate. Specifically, we assume that if the speaker knew that Billy ate pizza, then he would have uttered the alternative statement in (4), “Billy ate pizza”, whereas if he knew that Billy ate pasta, he would have uttered the statement in (5), “Billy ate pasta.” Since the speaker did not choose either of these stronger and more informative utterances, we conclude that the speaker either lacked evidence regarding their truth (i.e., Step III-a above) or believed them to be false (i.e., Step III-b). However, if the speaker believed that either statement was false, then it is unlikely that he would have made the original statement in (3). 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 ate pizza, or whether or not Billy ate pasta – all he knows is that Billy ate one of these two items (for discussion, see Fox, 2007; Sauerland, 2004; Spector, 2003).[2]

Critically, by Neo-Gricean accounts, ignorance implicature – i.e., Step IIIa above – is a necessary component of the stronger, scalar implicature in Step IIIb. In order to arrive at the inference that a speaker believes a stronger alternative statement to be false, the listener must first be able to retrieve from the original utterance the implication that it's not the case that the speaker believes the stronger alternative to be true. Only then, on this view, can the listener make the assumption that the speaker has knowledge regarding the stronger statements they didn't make, and in particular that they believe them to be false.

Given this framework, it becomes possible to probe the specific epistemic abilities that are potentially relevant to scalar implicature, rather than assuming that individuals who have *any* pragmatic deficit must therefore lack the epistemic resources required by implicature. If individuals with ASD compute scalar implicatures and do so via Gricean epistemic reasoning, then they should also successfully compute ignorance implicatures, since it is a necessary part of scalar implicature on the neo-Gricean hypothesis. On the other hand, evidence that individuals with ASD compute scalar implicatures despite an inability to infer ignorance, would lend support

[2] There is an important distinction between an ignorance implicature, the inference that the speaker doesn't know whether a given proposition is true or false, versus a weak inference, the inference that the speaker doesn't believe a position be true but it is unclear whether the speaker is ignorant (see the discussion in Geurts, 2010). However, this distinction is not critical in relation to the experimental results presented in this paper. Therefore, we will call both type of inference ignorance implicatures for the sake of the following discussion.

to the idea that implicatures can arise via non-Gricean mechanisms – e.g., by default whenever a relevant scalar term is used, or via a non-epistemic grammatical operator (Levinson, Chierchia et al., Fox, 2007).

A limitation to this logic, however, is that although failure to compute ignorance implicatures would support relatively strong conclusions, success is harder to interpret. This is because an ability to compute ignorance implicatures would be consistent both with a situation in which scalar implicature involves epistemic reasoning, and one in which it doesn't – i.e., where ignorance implicatures are computed via epistemic reasoning, but scalar implicatures via a non-epistemic computation, like a grammatical operator (e.g., Chierchia et al., 2012; Fox, 2007). For this reason, we also administered a second task, in which we asked whether subjects suspend scalar implicatures based on the epistemic states of the speaker (for a similar task, see Bergen & Grodner, 2012; Goodman & Stuhlmueller, 2013). As mentioned previously, on standard Neo-Gricean accounts, scalar implicature involves an assumption on the part of the listener that the speaker is knowledgeable regarding the truth of the stronger statement (i.e., the "epistemic step"). Thus, when this assumption is not warranted (e.g., the listener knows that the speaker does not have full knowledge about a given situation), listeners should not compute scalar implicatures. To illustrate this point, consider the sentence in (1), rewritten here as (6):

(6) Some of the students passed the exam.

Imagine, first, that this statement is uttered by a professor who has graded all of the exams in the class. This professor knows exactly how many students passed; thus, she chooses to utter the statement in (6) because she knows that the stronger alternative statement, "All of the students passed the exam" is false. Now, consider a context in which the statement in (6) is uttered instead by someone who only talked to several students in the class, each of whom happened to have passed the exam. This person does not know whether all of the students passed; she only knows that a few did. She therefore chooses the utterance in (6) because she does not know whether the stronger statement containing *all* is true. The scalar implicature that (6) implies "not all of the students passed the exam" is therefore not warranted in this case. Based on this logic, we presented subjects with statements in contexts where the speaker was clearly in a position to know whether a stronger alternative statement was true and in contexts where the speaker did not have enough evidence to know whether a stronger alternative statement was true. This allowed us to determine whether subjects reasoned about the speaker's knowledge state when deciding whether or not to compute scalar implicatures.

In summary, using two tasks we explored the role of epistemic reasoning in conversational implicature by testing high-functioning adolescents with ASD (aged 12 to 18 years). Building on methods developed in recent studies of ignorance implicature in typically developing children (Hochstein, Bale, Fox, & Barner, 2014), we asked (1) whether adolescents with ASD compute scalar implicatures in communicative contexts (2) whether they are able to compute ignorance implicatures – i.e., inferences commonly thought to underlie scalar implicature on Gricean accounts, and (3) whether they indeed reason about the epistemic states of interlocutors when computing implicatures in a task that manipulates whether speakers have partial or full knowledge.

EXPERIMENT

Methods

Participants

We tested 17 neurotypical adults (8 males; *M* = 22.6; range = 18-41) to establish the validity of our methods. In addition, we tested 18 adolescents with an autism spectrum disorder

(13 males, *M* = 14.9, range = 12-18). All subjects reported English as their first language. As we report below, it was not necessary to collect data for age, language, or IQ matched neurotypical adolescents, since the adolescents with ASD did not differ even from our neurotypical adults on any measures in the Ignorance Implicature or Scalar Implicature tasks, and only differed on 2 out of 4 critical trials in the Speaker Knowledge task (and did not differ on control trials in this task). Specifically, they differ on trials that required suspending scalar implicatures based on epistemic reasoning. In sum, a comparison to neurotypical adults makes a very strong case that adolescents with ASD are very much unimpaired in almost all cases.

The participants with ASD were all diagnosed by a licensed clinical psychologist or medical doctor not associated with this research, based on DSM-IV-TR or DSM-V-TR criteria (APA, 2004; 2013) and this diagnosis was confirmed using the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000). The ADOS (Lord et al., 2000) is a well-established measure of social and communicative behaviors that is used to diagnose and assess ASD. Subjects in this study were administered either Module 3 or Module 4 of the ADOS, depending on their age and maturity level, by a research assistant from an Autism laboratory at UCSD who had extensive experience administering the ADOS.

Intelligence was assessed using the Wechsler Intelligence Scale for Children (WISC) or the Wechsler Abbreviated Scale of Intelligence (WASI; Weschler, 1939). Exclusionary criteria included a full-scale IQ score below 70. Three additional subjects were excluded because their IQ scores (67, 52, and 64) fell below this cutoff.

Finally, none of the subjects included in this study had any other known neurological or genetic conditions or significant hearing, visual, or physical impairments. One additional subject was excluded because he was diagnosed with a rare genetic and neurological condition.

Design

Subjects were tested on four different tasks: (1) an *Ignorance Implicature* task, (2) a *Scalar Implicature* task (3) a *Partial Knowledge* task, and (4) a *2-scales* task (which compared disjunction to *some*/*all* implicatures in an act-out paradigm, and is not reported in this study). The first two tasks were counterbalanced in order between subjects. The third and fourth tasks always occurred in third and fourth position, respectively. Once these tasks were completed, subjects were also administered the Peabody Picture Vocabulary test (PPVT; Dunn & Dunn, 1997), a measure of vocabulary size, and the Strange Stories Test (Happé, 1994), a measure of Theory of Mind.

Ignorance Implicature task: This task was adapted from Hochstein, Bale, Fox, and Barner (2014). In the 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 four 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 character. Subjects were given corrective feedback if they answered a question incorrectly on the warm-up trials.

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 (with 2 trials for each type), as shown in Table 1, and these trials were presented in one of two counter-balanced orders.

Table 1: Test trials in the Ignorance Implicature Task

Condition	Choices	Animal Takes	Someone Says…	Correct
True-1 (control)	A / B	B	The bear took B.	Seeing doll
True-2 (control)	A / B	A & B	The bear took A & B	Seeing doll
False (control)	A / B	B	The bear took A.	Blindfolded doll
Ign. 1	A / B	B	The bear took A or B.	Blindfolded doll
Ign. 2	A / B	A & B	The bear took A or B	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. 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. 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 subjects could compute scalar implicatures but not 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 if subjects attributed *Ignorance-1* statements to the blindfolded puppet it was not simply because they noticed a discrepancy between the number of items taken (i.e., one) and the number of items mentioned (i.e., two) on these trials.

Scalar Implicature Task: This task closely resembled the *Ignorance Implicature* task, but differed crucially in that subjects attributed sentences to smart vs. silly speakers rather than to knowledgeable vs. ignorant ones. 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. Again, subjects were given corrective feedback if they answered a question incorrectly.

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 said that?"). At the end of each trial, subjects were asked to justify their choices. These test trials consisted of 4 different types (with 2 trials for each type), as shown in Table 2, and these trials were presented in one of two counter-balanced orders.

Both of the *True* control trials were attributable to the Smart Puppet and the *Underinformative* control trials to the *Silly Puppet.* For the underinformative control trials we used numbers, since typically-developing children 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.*

Table 2: Test trials in the Scalar Implicature Task

Condition	Bear has	Cow has	Someone says…	Correct
True-1 (control)	A	B	Each animal has A or B	Smart Puppet
True-2 (control)	A & B	A & B	Each animal has A & B	Smart Puppet
Number (control)	3 A	3 A	Each animal has 2 A.	Silly Puppet
Scalar OR	A & B	A & B	Each animal has A or B	Silly Puppet

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 the universal quantifier, disjunctive statements typically generate an ignorance implicature, whereas disjunctive statements containing "each" usually do not. 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 subjects' 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.

Partial Knowledge Task: In this task, inspired by Bergen and Grodner (2012), the experimenter put three boxes on the table and introduced subjects to a plastic figurine named Farmer Brown. Subjects were then told, "Farmer Brown doesn't know what's inside the boxes, but he's going to look – sometimes in all of them and sometimes only in 2 of them. Then he's going to tell you what he knows. He's trying to help you; he's not trying to trick you. Then, I'm going to ask you some questions about the boxes, and you can use what you know and what Farmer Brown tells you to say, 'yes' or 'no.' But sometimes you might not be able to know what's inside a box. Then you can just say, 'I don't know.' So your options are 'yes,' 'no,' and 'I don't know.'"

On each trial, Farmer Brown looked inside a certain number of boxes and then made a statement about their contents, as described below. The experimenter then asked the subject a question about the third box. The trial types varied on two dimensions: the speaker's knowledge (*Full-knowledge* vs. *Partial-knowledge*) and the term used by the speaker (*all*, *some*, or *two*). There were 5 total trial types, as described in Table 3, and 3 trials for each trial type. Farmer Brown always opened the first and second boxes such that their contents were fully visible to the subject. These boxes always had two pieces of the same fruit inside (i.e., each box had two bananas, two oranges, two grapes, or two strawberries inside). On *Full-knowledge* trials, Farmer Brown also peeked inside the third box such that he could see what was inside but the subject could not. On *Partial-knowledge* trials, Farmer Brown did not look inside the third box at all. Once Farmer Brown had finished looking inside the boxes, the experimenter pointed to the third box and asked subjects, "Does Farmer Brown know what's in this box?"[3] Once subjects

[3] Subjects were only allowed to proceed once they responded to this question correctly. If subjects answered incorrectly, the experimenter demonstrated again which boxes Farmer Brown had looked inside, and asked subjects the question again.

responded correctly, the experimenter said, "Now Farmer Brown is going to tell you what he knows." On *Some* trials, Farmer Brown said, "Some of the boxes have [strawberries]" and on *Two* trials, Farmer Brown said, "Two of the boxes have [strawberries]." On *all* trials (which served as controls and only occurred with *Full-knowledge*), Farmer Brown said, "All of the boxes have [strawberries]." The experimenter then pointed to the third box and asked subjects, "Do you think there are [strawberries] in this box?" Critically, although we directly asked subjects whether the speaker, Farmer Brown, was knowledgeable, this was not our primary measure of interest. Instead, we were interested in whether subjects *used* this information in the service of scalar implicature, or instead ignored the speaker's mental states.

Table 3: Partial Knowledge Task trial types

Knowledge	Description	Farmer Brown said…	Correct re: Box 3
Full-knowledge	*All*	"All of the boxes have *x*"	'Yes'
	Some	"Some of the boxes have *x*"	'No'
	Two	"Two of the boxes have *x*"	'No'
Partial-knowledge	*All*	----	----
	Some	"Some of the boxes have *x*"	'I don't know'
	Two	"Two of the boxes have *x*"	'I don't know'

On *Full-knowledge* trials, when Farmer Brown uttered a statement like, "Some of the boxes have strawberries", subjects were clearly licensed to compute the scalar implicature that not all of the boxes contained strawberries. In contrast, when Farmer Brown uttered these same statements on the *Partial-knowledge* trials, subjects were not licensed to make the implicature, as Farmer Brown was not in a position to know whether all of the boxes had strawberries. Thus, if subjects reasoned about Farmer Brown's epistemic state when deciding whether or not to compute a scalar implicature, they should have computed a scalar implicature only on the *Full-knowledge* trials.

On the *Partial-knowledge* trials, subjects had no information about the contents of the third box, and so they should have answered, “I don’t know” to the critical question (or, randomly chosen between “Yes” and “No”).

Note that the *Full-knowledge some* trials in this task required subjects to compute scalar implicatures in a way that resembles how these inferences are computed in everyday conversation – i.e., by inferring something about the world based on speaker’s choice of a weaker utterance over a stronger alternative one. In most previous tests of scalar implicature (including the first task described above), subjects do not actually have to do this; instead, they are provided with a true state of affairs in the world and are simply asked to judge the appropriateness of an utterance in describing this state of affairs (or, in our experiment, who likely uttered the sentence). Subjects’ performance on the *full-knowledge some* trials in this task may therefore provide a more naturalistic assessment of their ability to compute scalar implicatures than previous studies (e.g., Pijnacker et al., 2009; Chevallier et al., 2010) and even the *Scalar Implicature* task described above.

Peabody Picture Vocabulary Test (*PPVT*)*:* The PPVT (Dunn & Dunn, 1997) is a well-established measure of receptive vocabulary that is widely used in assessments of children’s language skills. Subjects are presented with a word and are asked to identify which of four pictures this word refers to.

Strange Stories Test: This task was designed by Happé (1994) to assess Theory of Mind abilities in older subjects by presenting them with short stories about “everyday situations where people say things they do not literally mean” and asking them to explain the characters’ intentions (130). For instance, in one story, a girl playfully holds up a banana to her ear and says, “This banana is a telephone.” In another story, a girl breaks her mother’s favorite vase and then tells her mother "the dog knocked it over; it wasn't my fault!" For each story, subjects are asked, "Is it true what [the character] said?" and "Why did he/she say that?" Subjects' responses were scored on two separate dimensions: 1) whether they made reference to mental states (coded as proportion mental state use), and 2) whether they were correct or incorrect (coded as proportion correct).

2.2 Results

Prior to analysis, we compared performance on the two different types of control trials within each task and found no significant differences between them. 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 within each task (but not across tasks). Figure 1 displays performance by ASD subjects and neurotypical adults on the critical and control trials.

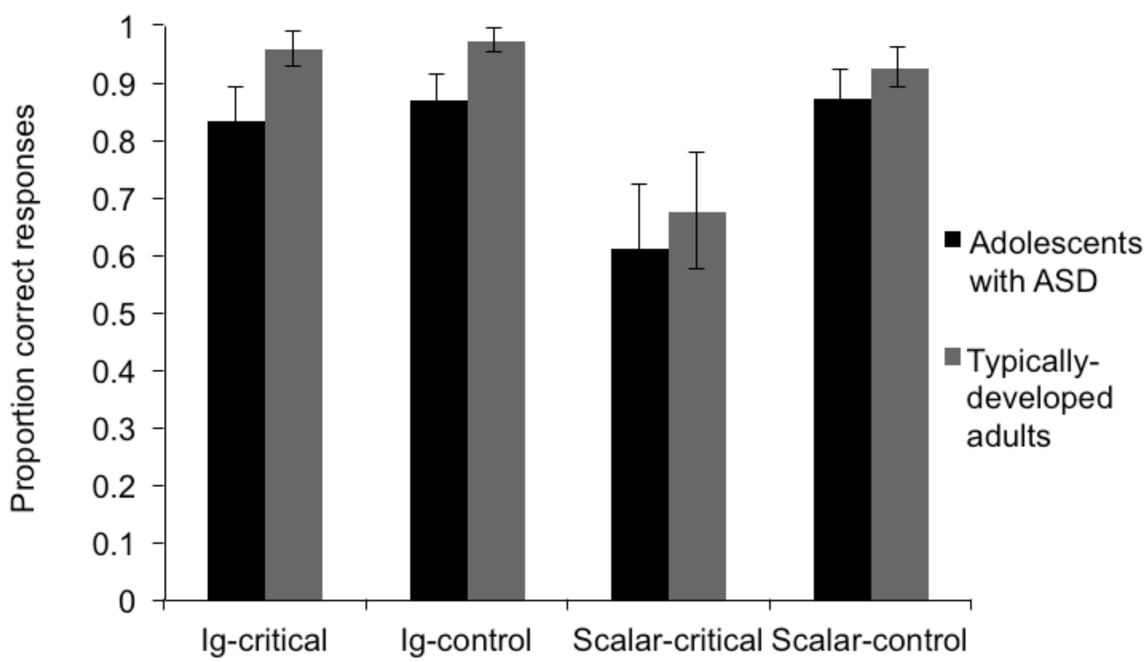

Figure 1: Proportion of correct responses by ASD subjects and neurotypical adults on critical and control trials in the Ignorance and Scalar Implicature tasks. Error bars represent standard error of the mean.

First, we compared our ASD subjects' performance on the *Ignorance Implicature* task to that of neurotypical adults. We used a binomial logit mixed effects model to predict binary response (correct vs. incorrect) using Group (ASD vs. Typical) and Trial Type (Critical vs. Control) as fixed variables and subject as a random factor. This model found no significant effects, suggesting that the adolescents with ASD performed no differently from neurotypical adults on the *Ignorance Implicature* task.

Next, we compared our ASD subjects' performance on the *Scalar Implicature* task to that of neurotypical adults. Again, we used a binomial logit mixed effects model to predict binary response (correct vs. incorrect) using Group (ASD vs. Typical) and Trial Type (Critical vs. Control) as fixed variables and subject as a random factor. This model found a main effect of trial type (Odds ratio = 0.1281, z = -3. 451, p = .00056) suggesting that subjects performed better overall on control trials than on critical trials in this task. However, there were no other significant effects. Thus, the adolescents with ASD correctly attributed disjunctive trials involving scalar implicature to the "silly puppet" to the same degree as neurotypical adults.

We next compared performance across the scalar implicature and ignorance implicature tasks within each group. For each group we used a binomial logit mixed effects model to predict binary response (correct vs. incorrect) using Task (Ignorance Implicature vs. Scalar Implicature) and Trial Type (Critical vs. Control) as fixed variables and subject as a random factor. For the ASD group, the model found an interaction of Task and Trial Type (Odds ratio = 0.1649, z = -2.27, p = .023), but no other significant effects. This interaction was driven by a significant difference between critical (*M* = .61, *SE* = .11) and control (*M* = .87, *SE* = .05) trials in the *Scalar Implicature* task (Wilcoxon T = 7, N = 10, *p* = .039) but no significant difference between critical (*M* = .83, *SE* = .06) and control (*M* = .87, *SE* = .04) trials in the *Ignorance Implicature* task (Wilcoxon T = 14, N = 8, p > .05). In the neurotypical adult group, the model found a marginally significant interaction of Task and Trial Type (Odds ratio = 0.1692, z = -1.66, p = .097). Thus, both groups easily computed ignorance implicatures, and in fact showed a smaller difference between critical and control trials in this task than they did on the scalar implicature task.

Table 4: Age, IQ, and PPVT scores for the ASD group in Experiment 1

	Mean	**Range**	**St. Dev.**
Age	14.9	12-18	1.9
IQ	97	71-132	18.9
PPVT[4]	101	56-130	22.2

In the ASD group, we next explored whether there was any relation between performance on the two implicature tasks and age, IQ, vocabulary size (as measured by the PPVT), or Theory of Mind (as measured by the Strange Stories task). See Table 4 for descriptive statistics. With the exception of Age, all of the variables were very highly correlated with each other, as indicated by Pearson's correlation (see Table 5). We therefore performed forward and backward stepwise model selection to determine the best fitting model to predict performance on the Ignorance Implicature task. Based on this, we used a mixed effects model to predict response using only Trial Type and PPVT as variables. This model found a main effect of PPVT (beta = 0.468, z = 2.26, p = 0.032) but no other significant effects. We next performed forward and backward stepwise model selection to determine the best fitting model to predict performance on the Scalar Implicature task. Based on this, we used a mixed effects model to predict response using only Trial Type, WASI, and Age as variables. This model found a main effect of Trial Type (beta = -2.91, z = -2.340, p = 0.027) and a marginally significant interaction of Trial Type and Age (beta = 5.76, z = 1.90, p = 0.068) but no other significant effects.

[4] One subject did not complete the PPVT. This subject is therefore excluded from any analyses regarding PPVT scores.

Table 5: Pearson's Correlations for WASI, PPVT, Strange Stories, and Age

	1	2	3	4	5
1. WASI	-				
2. PPVT	0.91	-			
3. Strange Stories 1[1]	.74	.78	-		
4. Strange Stories 2[2]	.83	.91	.89	-	
5. Age	.22	.16	.20	.14	-

[1] Proportion Mental State
[2] Proportion Correct

Partial Knowledge task

Figure 2 displays performance by ASD subjects and neurotypical adults on the critical and control trials in both tasks. Since both groups performed very well on the control trials in this task, we did not include these trials in subsequent analyses. Instead, we contrasted performance on Full-Knowledge and Partial-Knowledge trials, thereby allowing us to directly test whether these trial types differed.

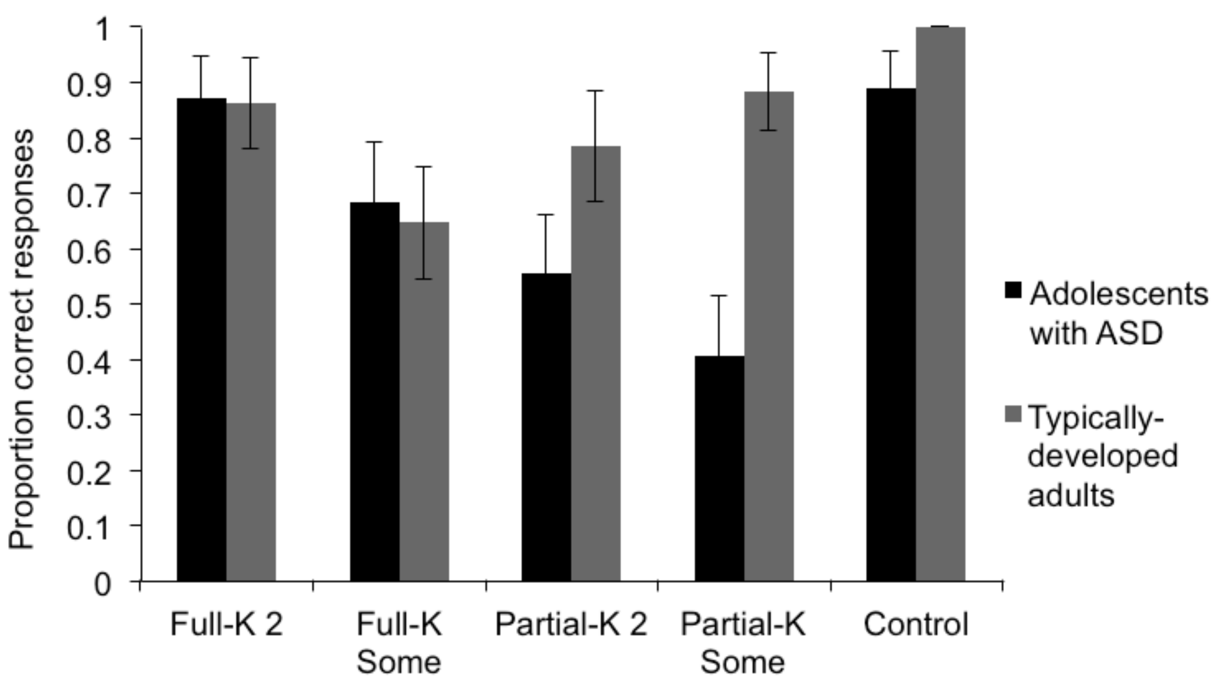

Figure 2: Proportion of correct responses by ASD subjects and neurotypical adults on the Partial Knowledge task. Error bars represent standard error of the mean.

We first analyzed performance on the *Partial Knowledge* task in a model that included data from both individuals with ASD and neurotypical adults. We used a binomial logit mixed effects model to predict binary response (correct vs. incorrect) using Group (ASD vs. Typical), Condition (Full-Knowledge vs. Partial-Knowledge), and Term (Some vs. Two) as fixed variables and subject as a random factor. This model found a marginally significant effect of Term (Odd ratio = 8.63, z = 1.86, $p = .063$), suggesting that subjects performed better overall on "two" trials than on "some" trials (this effect was also found in both groups when data were analyzed separately; for the ASD group, there was a marginally significant effect of Term (Odds ratio = 11.88, z = 1.89, $p = .0589$) and for the neurotypical adult group there was a highly significant effect of Term (Odds ratio 5.05, = $z = 2.847$, $p = 0.0056$). In addition, this model found an interaction between Group and Condition (Odds ratio = 21.83637 $z = 2.763$, $p = .0057$). This was driven by a significant difference between Full-Knowledge ($M = .78$, $SE = .066$) and Partial-Knowledge trials ($M = .48$, $SE = .074$) in the ASD group (Wilcoxon T = 0, N = 11, $p = .0036$), but no significant difference between Full-Knowledge ($M = .75$, $SE = .076$) and Partial-Knowledge trials ($M = .83$, $SE = .068$) in the neurotypical adult group (Wilcoxon T = 24, N = 11, $p = .43$). Post hoc tests found that when individuals with ASD provided incorrect responses on Partial Knowledge trials, they were significantly more likely to respond "no" than "yes" consistent with computing a scalar implicature. This was true both for "two" trials (86.4% "no" responses; Wilcoxon T = 8.5, N = 13, p = .01) and for "some" responses (78.8% "no" responses; Wilcoxon T = 12.5, N = 13, p = .022). Interestingly, on the small proportion of trials on which adults made incorrect responses, they always provided "no" responses, suggesting that when they erred it was also in the direction of infelicitously computing scalar implicatures, without epistemic justification. In sum, ASD subjects exhibited a somewhat diminished sensitivity to speaker knowledge when evaluating statements, despite computing implicatures to the same degree as neurotypical adults on Full Knowledge trials (for *some* and *two*), and despite similar performance on control trials, and when they made errors were highly likely to compute scalar implicatures without epistemic justification.

In the ASD group, we next explored whether there was any relation between age, IQ scores, PPVT scores, and Strange Stories test scores and performance on Full-Knowledge vs. Partial-Knowledge trials. Since the IQ, PPVT, and Strange Stories test scores were highly correlated we performed forward and backward stepwise model selection to determine the best fitting model. Based on this, we used a mixed effects model to predict response using only Condition, Age, and IQ as variables. This model found a main effect of IQ (beta = 0.513, z = 3.124, p = 0.00412) but no other significant effects.

Taken together, these findings suggest that although the ASD group correctly inferred that statements like, “Some of the boxes have strawberries” and “Two of the boxes have strawberries” implied that not *all* of the boxes have strawberries on Full-Knowledge trials, and did so to the same degree as neurotypical adults, they did not consistently cancel these inferences on Partial-Knowledge trials.

GENERAL DISCUSSION

This study explored implicature computation in adolescents with ASD to determine, first, whether they can compute scalar implicatures, second, whether they are capable of the epistemic reasoning that neo-Gricean theories argue are constitutive of implicature, and, third, whether they actually engage in this kind of epistemic reasoning when computing scalar implicatures.

This study revealed three major findings. First, we found that high-functioning adolescents with ASD compute scalar implicatures to the same degree as neurotypical adults. Second, we found that there was also no difference between these groups in their ability to compute ignorance implicatures, which on neo-Gricean theories are required for scalar implicature. Third, we found that even though adolescents with ASD showed awareness of speakers' mental states when asked directly what a speaker knew or did not know, they often failed to spontaneously consider speakers' specific epistemic states when actually computing scalar implicatures. In the Partial Knowledge task, subjects with ASD computed scalar implicatures both for fully knowledgeable speakers and for speakers that were only partially knowledgeable.

Together, the results from this study suggest that many high-functioning individuals with ASD perform quite well on tests of scalar implicature. These results are consistent with those of previous reports that investigated adolescents with ASD (Pijnacker et al., 2009; Chevallier et al., 2010), and extend these results to experimental conditions that involve communicative exchanges that more closely resemble how implicatures are computed in conversation. However, our data also suggest that although adolescents with ASD are capable of reasoning about epistemic states when explicitly required to, they struggle to suspend scalar implicatures when they should be precluded by the ignorance of a speaker – i.e., a speaker who had partial knowledge.

At first pass, this collection of results is puzzling: How might high-functioning adolescents with ASD compute ignorance implicatures despite failing to consider the epistemic states of speakers when computing scalar implicatures? To help understand these results, it is instructive to again consider the steps thought to be involved in scalar implicature according to different accounts. These are repeated from the Introduction, below:

I. Compute the literal meaning of a sentence S

e.g., Some of the students passed the exam

II. Generate a set of relevant scalar alternatives to S

e.g., All of the students passed the exam

III. Strengthen the meaning of S by negating the relevant stronger alternatives.[5]

e.g., Some, but not all, of the students passed the exam

In the Introduction, we noted that on neo-Gricean accounts this third step – i.e., negating stronger alternatives – can be broken down into two sub-steps, each of which has an epistemic component not present in the previous two steps. These are described as follows:

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

IIIb. Given evidence that the speaker is knowledgeable, infer that the speaker believes *p* to be false – i.e., $B(\neg p)$).

Thus, on neo-Gricean accounts, whenever a scalar implicature is computed, it should arise as the product of epistemic reasoning. Although non-Gricean accounts, like the grammatical view, do not posit epistemic reasoning as part of implicature (instead arguing that ignorance implicature is an independent form of inference), all theories predict that neurotypical individuals will consider a speaker's epistemic states as a precondition to computing implicatures, whether this epistemic reasoning is built into the implicature algorithm (as on the Gricean view) or is a separate computation (as on the grammatical view). One possible source of this difference is that only the *Ignorance Implicature* task requires subjects to base judgments on

[5] Technically, not just "stronger" alternatives, but also those that are not weaker than the original sentence.

the knowledge states of speakers, whereas the *Partial Knowledge* task does not explicitly demand such reasoning. In the *Ignorance Implicature* task, one candidate speaker is clearly knowledgeable and the other is clearly ignorant, and subjects are asked to attribute utterances to one or the other speaker based purely on this difference. However, in the *Partial Knowledge* task, although the knowledge state of the speaker regarding the contents of a box is manipulated across trials and even explicitly queried as part of the task, the dependent variable is never explicitly related to the speaker's knowledge. Instead, subjects are asked what *they* think is contained in the box. Thus, in this case, the dependent measure is what the *subject* believes based on what was said, not what the speaker believed. In sum, high-functioning individuals with ASD may be able to reason about epistemic states when explicitly required to do so, but may not spontaneously consider speaker states when it is not required, even when these states are known to them as they were in our study. On this analysis, adolescents with ASD may not consider speaker states to compute scalar implicatures because these inferences are not obligatorily epistemic in nature.

These results may have important consequences for theories of implicature, but should also be interpreted with a degree of caution, in light of alternative interpretations. On face value, the data indicate that at least some individuals – who are able to make sophisticated inferences about the mental states of speakers – nevertheless do not spontaneously use this mental state information to modulate their computation of implicatures. This suggests that, at least in this group, scalar implicatures can be computed without reference to mental state information. This finding can be interpreted in a number of ways. One possibility is that individuals with ASD compute implicatures identically to neurotypical adults – i.e., via a non-epistemic, grammatical mechanism – and that the two groups only differ in that neurotypical adults can condition these inferences on computationally independent epistemic factors. On the grammatical view, implicatures can be suspended if a speaker is known to be ignorant, but when implicatures *are* computed, their representational format is non-epistemic, and does not involve ignorance implicature. Another possibility is that different populations compute implicatures in different ways. Whereas neurotypical adults may use a full-fledged neo-Gricean route, complete with epistemic content, individuals with ASD may deploy heuristic shortcuts that bypass epistemic reasoning, similar to Levinson (2000).

In addition to these theoretically interesting alternative accounts, there are also potentially less rich explanations our data. For example, in our study we did not compare individuals with ASD to subjects matched according to chronological or mental age. Instead, they were only compared to neurotypical college students. To the extent that our conclusions rely on this comparison, it is possible that non-epistemic differences between groups may have explained aspects of our findings. Although this question does not arise for the scalar implicature and ignorance implicature tasks, where no important differences were found between groups, it does arise for the Speaker Knowledge task, since here the adolescents with ASD performed differently. For example, test questions that involved using speaker knowledge to infer the contents of a covered box may have exceeded the working memory capacity of the adolescent group. While this is possible, it is important to note two rejoinders. First, there was no significant difference between adults and adolescents with ASD on control items, which were arguably as taxing the domain-general capacities as the critical test items. Only critical items that involved implicature differed. Second, the groups did not differ on the Ignorance Implicature task or the preliminary Scalar Implicature task. Both of these tasks were similarly complex and arguably involved comparable demands on working memory. Most important, the critical comparison in the Speaker Knowledge task was not that between adults and adolescents with ASD, but instead was that between critical and control trials within each group. Although a comparison to age-matched subjects could contribute to understanding the specificity of behaviors to individuals with ASD – e.g., if neurotypical teens performed similarly to college students – evidence that age-matched controls perform similarly would remain difficult to interpret, since this result, while consistent with an age-related problem with working memory, would not rule out the possibility that neurotypical adolescents also fail to spontaneously reason about epistemic states in the service of implicature.

Beyond the possible implications of this work for theories of implicature, our study – and previous work on young typically developing children – suggests that caution should be applied when using either population as a model for global pragmatic deficit, as some previous developmental studies of language development have done. As noted in the Introduction, past studies have treated neurotypical preschoolers and individuals with ASD as case studies for examining the effects of global pragmatic deficits, based on the assumption that whatever abilities these groups have, they must not be supported by pragmatic inference. For example, in a recent paper by de Marchena et al. (2011), ASD was used as a test case for exploring the role of pragmatic reasoning in word learning. According to these researchers, the fact that children with ASD acquire large vocabularies despite their pragmatic deficits suggests that vocabulary acquisition does not depend on pragmatic reasoning. However, as we have seen, there are dissociations between different forms of epistemic reasoning and performance on different pragmatic tasks. Thus, it may be premature to assume that individuals who are impaired on one type of pragmatic task should be impaired on all others, and thus that they are a good model for generalized pragmatic deficit. Similarly, we cannot assume that if these groups succeed at a task ordinarily thought to involve pragmatic reasoning then the task must not rely on pragmatics after all. Instead, given the somewhat non-monolithic nature of pragmatics, it may be necessary to directly test the specific form of epistemic inference of interest and how this epistemic reasoning is related to word learning or conversational inference.

To conclude, the findings from this study suggest that although individuals with ASD exhibit deficits with theory of mind and certain pragmatic aspects of language, these deficits are not strong enough to preclude the types of epistemic reasoning required for Gricean conversational implicatures, like ignorance implicature. High-functioning adolescents with ASD appear able to reason about the knowledge states of interlocutors when explicitly asked to do so. However, our findings also suggest that individuals with ASD may not spontaneously reason about the knowledge states of specific speakers when computing scalar implicatures, suggesting that scalar implicature computation may not strictly require Gricean epistemic reasoning.

Acknowledgments

Many thanks to Danny Fox, Karen Dobkins, Melissa McIntire, Shereen Cohen, Katherine Stavropoulos, Katie Kimura, Katie Wagner, Laura Taouk, Troy Campbell, Daisy Hong, Danny Fox, Valerie's List, 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 grant from the James S. McDonnell foundation.

REFERENCES

American Psychiatric Association. (2004). Diagnostic and statistical manual of mental disorders (4th-TR ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th-TR ed.). Washington, DC: American Psychiatric Association.

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child development*, 62, 874-890.

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

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

Baron-Cohen, S. (1988). Social and pragmatic deficits in autism: Cognitive or affective? *Journal of autism and developmental disorders*, 18, 379-402.

Baron-Cohen, S. (1995). Mindblindness: An essay on autism and theory of mind. Cambridge, MA: MIT Press

Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a "theory of mind"?. *Cognition*, 21, 37-46.

Bergen, L., & Grodner, D. J. (2012). Speaker knowledge influences the comprehension of pragmatic inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 1450-1460.

Chevallier, C., Wilson, D., Happé, F., & Noveck, I. (2010). Scalar inferences in autism spectrum disorders. *Journal of autism and developmental disorders*, 40, 1104-1117.

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

Chierchia, G., Fox, D., & Spector, B. (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. New York: Mouton de Gruyter.

Clark, E. (1990). On the pragmatics of contrast. *Journal of Child Language*, 17, 417–431.

de Marchena, A., Eigsti, I. M., Worek, A., Ono, K. E., & Snedeker, J. (2011). Mutual exclusivity in autism spectrum disorders: Testing the pragmatic hypothesis. *Cognition*, 119, 96-113.

Dalrymple, M., Kanazawa, M., Kim, Y., Mchombo, S., & Peters, S. (1998). Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy* 21: 159-210.

Diesendruck, G. & Markson, L. (2001). Children's Avoidance of Lexical Overlap: A Pragmatic Account. *Developmental Psychology,* 37, 630-641.

Dunn, L. M., & Dunn, L. M. (1997). *Examiner's manual for the PPVT-III peabody picture vocabulary test*. AGS.

Eales, M. J. (1993). Pragmatic impairments in adults with childhood diagnoses of autism or developmental receptive language disorder. *Journal of Autism and Developmental Disorders*, 23, 593-617.

Eisenmajer, R., Prior, M., Leekam, S., Wing, L., Ong, B., Gould, J., & Welham, M. (1998). Delayed language onset as a predictor of clinical symptoms in pervasive developmental disorders. *Journal of autism and developmental disorders*, 28, 527-533.

Foppolo, F., Guasti, M.T. & Chierchia, G. (2012). Scalar implicatures in child language: Give children a chance. *Language Learning and Development*, 8, 365-394.

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

Frith, U. (2001). Mind blindness and the brain in autism. *Neuron*, 32, 969–979

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

Geurts, B. (2010). Quantity Implicatures. Cambridge: Cambridge University Press.

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

Happé, F. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, 48, 101-119.

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of autism and Developmental disorders*, 24, 129-154.

Hochstein. L, Bale, A., Fox, D., & Barner, D. (2014). Ignorance and Inference: Do problems with Gricean epistemic reasoning explain children's difficulties with scalar implicature? *Journal of Semantics.*

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

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

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT press.

Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition*, 108, 732-739.

Lord, C., Risi, S., Lambrecht, L., Cook Jr, E. H., Leventhal, B. L., DiLavore, P. C. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of autism and developmental disorders*, 30, 205-223.

Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive psychology*, 47, 241-275.

Meyer, M-C. (2013). Ignorance and Grammar. Ph.D. Dissertation, Massachusetts Institute of Technology. Cambridge, Massachusetts.

Minshew, N. J., Goldstein, G., & Siegel, D. J. (1995). Speech and language in high-functioning autistic individuals. *Neuropsychology*, 9, 255.

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

Ozonoff, S., & Miller, J. N. (1996). An exploration of right-hemisphere contributions to the pragmatic impairments of autism. *Brain and Language*, 52, 411-434.

Papafragou, A., & Tantalou, N. (2004). Children's computation of implicatures.*Language Acquisition*, 12, 71-82.

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

Pijnacker, J., Hagoort, P., Buitelaar, J., Teunisse, J. P., & Geurts, B. (2009). Pragmatic inferences in high-functioning adults with autism and Asperger syndrome. *Journal of autism and developmental disorders*, 39, 607-618.

Sabbagh, M. A., & Baldwin, D. A. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child development*, 72, 1054-1070.

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

Shriberg, L. D., Paul, R., McSweeny, J. L., Klin, A., Cohen, D. J., & Volkmar, F. R. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *Journal of Speech, Language & Hearing Research*, 44, 1097-1115.

Skordos, D., & Papafragou, A. (to appear). Children’s derivation of scalar implicatures: alternatives and relevance. *Cognition*.

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

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

Stiller, A., Goodman, N. D., & Frank, M. C. (2011). Ad-hoc scalar implicature in adults and children. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society,* Boston, July.

Tager-Flusberg, H. (1999). A psychological approach to understanding the social and language impairments in autism. *International review of psychiatry*, 11, 325-334.

Tager-Flusberg, H., Paul, R., & Lord, C. E. (2005). Language and communication in autism. In F. Volkmar & R. Paul & A. Klin & D. J. Cohen (Eds.), Handbook of autism and pervasive developmental disorder, 1, 335-364, New York: Wiley.

Tomasello, M., & Barton, M. E. (1994). Learning words in nonostensive contexts. *Developmental psychology*, 30, 639-650.

Wechsler, D. (1939). The measurement of adult intelligence. Baltimore, MD: Williams & Witkins.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child development*, 72, 655-684.

Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind Tasks. *Child development*, 75, 523-541.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103-128.