UC San Diego
UC San Diego Electronic Theses and Dissertations

Title
Epistemic reasoning and implicature computation in typically-developing children and individuals with an Autism Spectrum Disorder /

Permalink
https://escholarship.org/uc/item/5bb3x5dg

Author
Hochstein, Lara Klainerman

Publication Date
2014

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Epistemic reasoning and implicature computation in typically-developing children and individuals with an Autism Spectrum Disorder

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Linguistics

by

Lara Klainerman Hochstein

Committee in charge:

Professor David Barner, Chair
Professor Andrew Kehler, Co-Chair
Professor Ivano Caponigro
Professor Victor Ferreira
Professor Rachel Mayberry

2014

Copyright

Lara Klainerman Hochstein, 2014

All rights reserved.

The dissertation of Lara Klainerman Hochstein is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____
Co-Chair

_____
Chair

University of California, San Diego

2014

TABLE OF CONTENTS

Signature page…… iii

Table of Contents……iv

List of Figures……...…...v

List of Tables……..vi

Acknowledgments…….vii

Vita…….....ix

Abstract of the Dissertation…….x

Chapter 1: Introduction……1

Chapter 2: Ignorance and Inference: do problems with Gricean epistemic reasoning explain children's difficulty with scalar implicature? …….....23

Chapter 3: Epistemic reasoning in children's early inferences: the case of ad hoc implicature……..66

Chapter 4: Scalar implicature in absence of epistemic reasoning? The case of Autism Spectrum Disorders …….. 92

Chapter 5: Conclusion…….....153

LIST OF FIGURES

Chapter 2:

Figure 2.1: Proportion of correct responses on critical and control trials in the Ignorance Implicature task.……43

Figure 2.2: Proportion of correct responses on critical and control trials in the Scalar Implicature task ……..44

Chapter 3:

Figure 3.1: Proportion of correct responses on critical vs. control trials in the Ad hoc and Ignorance Implicature tasks.……81

Chapter 4:

Figure 4.1: Proportion of correct responses by ASD subjects and typically-developing adults on critical and control trials in the Ignorance and Scalar Implicature tasks…115

Figure 4.2: Proportion of correct responses by ASD subjects and typically-developing adults on the Speaker Knowledge task …….....118

Figure 4.3: Proportion of correct responses on critical and control trials in the Ignorance and Scalar Implicature tasks by ASD age group.……..127

LIST OF TABLES

Chapter 2:

Table 2.1: Steps involved in Scalar Implicature computation……26

Table 2.2: Test trials in the Ignorance Implicature Task ……...38

Table 2.3: Test trials in the Scalar Implicature Task ……41

Chapter 3:

Table 3.1: Test trials in the Ignorance Implicature Task ……...79

Table 3.2: Test trials in the Ad hoc Implicature Task ……...80

Chapter 4:

Table 4.1: Description of Age, IQ scores, and PPVT scores for the ASD group in Experiment 1……105

Table 4.2: Test trials in the Ignorance Implicature Task……..108

Table 4.3: Test trials in the Scalar Implicature Task……111

Table 4.4: Test trials in the Speaker Knowledge Task…….113

Table 4.5: Description of Age, IQ scores, and PPVT scores for the ASD group in Experiment 2……126

ACKNOWLEDGMENTS

It is a truth universally acknowledged that the key to a happy and successful graduate career lies in the hands of one's advisor. I consider myself very fortunate to have had not one but two wonderful advisors by my side on this long and often challenging journey. First and foremost, I must thank Dave Barner for adopting me into his lab during my first year at UCSD and meeting with me almost every week since then – to help me design (or troubleshoot) my experiments, provide invaluable feedback on my papers, or just chat. He has been both a mentor and a friend, and without his constant encouragement and inspiration, I could never have made it this far. I am also incredibly grateful to Andy Kehler for his guidance on my research, for teaching me how to give clear and well-organized presentations, and for imparting his ever-increasing wisdom about the inner workings of the department to me.

Heartfelt thanks go out to the other members of my committee – Ivano Caponigro, Rachel Mayberry, and Vic Ferreira – for inspiring me in their classes and supporting me in my work. Thank you also to all the members of the Language and Development Lab, my second family in graduate school – but especially to Jess Sullivan, Katie Wagner, and Pierina Cheung for helping me with my statistical analyses and/or providing helpful feedback on my papers. I am equally indebted to Naja Ferjan Ramirez and Kate Davidson Zaremba (my "grad school guru") for their guidance, encouragement, and friendship over the past 6 years, and I owe a big caffeinated thank you to Monika Krach and Jess Thierman for keeping me company during long work sessions at various coffee shops around La Jolla. Many thanks are also due to Karen Dobkins, Melissa McIntire, and Katherine Stavropoulos for their crucial help with my ASD project, to Alan Bale and

Danny Fox for their contributions to the Ignorance paper, to the members of AndyLab and Developmental Brown Bag for listening to my research over and over again and always providing new and helpful feedback, to my research assistants Troy Campbell, Laura Taouk, and Daisy Hong for helping me run my studies, and to the many families that graciously donated their time to participate in my experiments.

In graduate school as in every other aspect of my life, my biggest fans and supporters have always been my family. Thank you to: Mom and Dad for, well, everything, to Ariel and Leor for their phone calls, snapchats, and visits, and to Ted and Michelle Hochstein for their support. Finally, thank you to my husband, Grant, for sharing my joys as well as my frustrations, for spoiling me during my final sprint towards the finish line, and for agreeing to get Lily, the most adorable stress reliever of all.

This work was supported by a fellowship from the Jacob K. Javits Foundation, awarded to the dissertation author, and a grant from the James S. McDonnell Foundation, awarded to David Barner.

****

Chapter 2, in full, has been submitted and may appear as a publication; Hochstein, L., Bale., A., Fox, D., and Barner, D. Ignorance and inference: do problems with Gricean epistemic reasoning explain children's difficulties with scalar implicature? Under review at the *Journal of Semantics*. The dissertation author was the primary investigator and author of this paper. Permission to reprint was granted by the *Journal of Semantics.*

Chapters 3 and 4, in full, are currently being prepared for submission for publication of all of the material presented here. The dissertation author was the primary investigator and author of this material.

VITA

2008	B.A. in French and Italian, Princeton University
2010	M.A. in Linguistics, University of California, San Diego
2014	Ph.D. in Linguistics, University of California, San Diego

ABSTRACT OF THE DISSERTATION

Epistemic reasoning and implicature computation in typically-developing children and individuals with an Autism Spectrum Disorder

by

Lara Klainerman Hochstein

Doctor of Philosophy in Linguistics

University of California, San Diego, 2014

Professor David Barner, Chair

Professor Andrew Kehler, Co-Chair

This dissertation explores the role of epistemic reasoning (i.e., reasoning about other people's knowledge, beliefs, and intentions) in implicature computation by addressing two seeming paradoxes: first, the fact that typically-developing children fail at specific inferences known as scalar implicature until relatively late in development despite exhibiting basic epistemic reasoning abilities from an early age, and, second, the

fact that individuals with Autism Spectrum Disorders (ASD) have been claimed to succeed at scalar implicature despite exhibiting deficits with epistemic reasoning in other domains.

Chapter 2 provides evidence that 5-year-olds successfully compute ignorance implicatures – inferences that involve significant epistemic reasoning about speaker knowledge and utterance informativeness – despite failing to compute scalar implicatures. On the basis of this finding, we argue that children's failures with scalar implicature do not stem from any difficulty with the epistemic reasoning involved and, instead, are most likely due to an inability to access the specific lexical alternatives involved in scalar implicature.

Results from Chapter 2 also show that 4-year-olds fail to compute both ignorance and scalar implicature, suggesting that the ability to compute basic Gricean inferences in language emerges around 5 years of age. This finding is somewhat at odds with claims in the literature that children can compute other forms of pragmatic inference such as *ad hoc* implicatures at an earlier age. Chapter 3 therefore explores the role of epistemic reasoning in children's ability to compute *ad hoc* implicatures. We show that 4-year-olds successfully compute *ad hoc* implicatures despite failing to compute ignorance implicatures, which raises the possibility that children compute *ad hoc* implicatures before they are able to engage in epistemic reasoning about speaker knowledge and utterance informativeness.

Finally, chapter 4 tests epistemic reasoning abilities in high-functioning children and adolescents with ASD. Results show that high-functioning adolescents with ASD successfully compute both scalar and ignorance implicature, while younger children with

ASD fail at both inferences. These results indicate that high-functioning individuals with ASD are capable of the kind of epistemic reasoning required to make basic inferences in language despite their other pragmatic deficits, although this ability nevertheless appears to be somewhat delayed in ASD.

CHAPTER 1:

Introduction

One of the most long-standing and spirited debates in the study of meaning regards the division of labor between semantics and pragmatics. Whereas the field of semantics has been concerned primarily with characterizing the literal meaning of utterances by determining context-independent conditions under which a sentence is true, pragmatics has been concerned primarily with bridging the gap between the literal, semantic meaning of utterances and the way these utterances are actually used and understood by speakers. Yet, although this distinction appears clear enough in theory, in practice it has often proven difficult to determine what is part of the semantic meaning of an utterance and what is part of its pragmatic meaning.

A key factor in distinguishing between semantic and pragmatic meaning is the role of epistemic reasoning – that is, reasoning about a speaker's knowledge, beliefs, and intentions. This kind of reasoning is not involved in the literal, semantic meaning of a word or utterance, but it is often essential for determining the intended, pragmatic meaning. Most research on epistemic reasoning has focused on the larger notion of Theory of Mind – generally defined as the ability to ascribe mental states to oneself and to other people, and to understand that other people's mental states can differ from one's own – and has explored the development of this ability in typically-developing children, special populations, and non-human primates, as well as the interface between Theory of Mind and language (e.g., Wellman et al., 2001; Russell et al., 1998; Baron-Cohen et al., 1985; Premack & Woodruff, 1978; de Villiers, 2007). However, researchers interested in

the interaction between semantics and pragmatics have also focused on the role of epistemic reasoning in various linguistic phenomena.

One such phenomenon that has been the focus of extensive research in recent years is a specific inference in language known as a scalar implicature. Scalar implicatures arise when a speaker utters a statement like, "Some of the buildings were destroyed in the fire" and listeners infer that *not all* of the buildings were destroyed. This inference has traditionally been considered a paradigmatic case of pragmatic enrichment via epistemic reasoning – i.e., by Gricean and Neo-Gricean accounts of implicature (e.g., Grice, 1975; Horn, 1972). Recently, however, it has been argued that scalar implicatures are computed via a grammatical process rather than via any reasoning about mental states, and thus may fall more under the domain of semantics than pragmatics (e.g., Chierchia et al., 2012).

To explore the role of epistemic reasoning in scalar implicature computation, researchers have turned increasingly to experimental study, 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 develops in children (e.g. Smith, 1980; Noveck, 2001; Papafragou & Musolino, 2003, among others) as well as in atypical populations, such as in individuals with an autism spectrum disorder (e.g., Pijnacker et al., 2009; Chevallier et al., 2010). Two major paradoxes regarding the role of epistemic reasoning in the computation of scalar implicature have emerged from this line of research, however. On the one hand, children have been shown to fail at scalar implicatures until relatively late in development, despite the fact that they exhibit basic epistemic reasoning abilities from

an early age. On the other hand, individuals with an Autism Spectrum Disorder (henceforth, ASD) have been claimed to succeed at scalar implicatures, despite the fact that they exhibit significant deficits with epistemic reasoning.

In this dissertation, we attempt to explain these two seeming paradoxes by disentangling the different factors at play in implicature computation. The first complicating factor is the fact that epistemic reasoning is non-monolithic and includes a wide range of abilities at different levels of complexity. Thus, it is possible for young children to have some early understanding of epistemic states yet still struggle with the kind of epistemic reasoning required to compute specific inferences in language, and it is similarly possible for individuals with ASD to fail at many pragmatic aspects of language yet still have sufficient epistemic reasoning ability to compute basic linguistic inferences. Our first aim in this dissertation, then, is to try to characterize more precisely what kind of epistemic reasoning is claimed to be involved in the computation of scalar implicature (i.e., by Gricean accounts) and to test more directly whether (and when) children and individuals with ASD are capable of this type of reasoning.

The second complicating factor is the fact that, by some accounts, scalar implicature does not actually require epistemic reasoning (e.g., Chierchia et al., 2012), and even by standard Gricean accounts, it does not *only* involve epistemic reasoning; it also involves some linguistic knowledge as well as significant processing resources (e.g., Barner et al., 2011; Pouscoulous et al., 2007). Thus, failure to compute scalar implicature does not necessarily reflect a difficulty with epistemic reasoning, and, conversely, success with scalar implicature does not necessarily reflect an ability to reason about epistemic states. Our second aim in this dissertation, then, is to pinpoint exactly which

aspects of scalar implicature computation pose problems for young children, and to investigate whether individuals with ASD actually engage in epistemic reasoning when computing these implicatures.

In exploring these issues, we hope to shed light not only on the development of epistemic reasoning abilities in typically-developing children and individuals with ASD but also on the nature of implicature more generally.

Scalar Implicatures

Paul Grice (1975) first proposed a theory of implicature in his seminal paper *Logic and Conversation*. According to Grice, conversation is conducted under the fundamental assumption that all parties involved share the same underlying communicative purposes. Without such a tacit agreement (i.e., if we could not be sure that our interlocutors would be cooperative), engaging in conversation would not be useful or worthwhile. Grice termed this mutual understanding the *Cooperative Principle*, and proposed that it is upheld primarily through adherence to four more specific maxims: the maxims of Quality, Quantity, Relation, and Manner. The Maxim of Quality specifies that speakers must communicate only that which they believe to be true and for which they have sufficient evidence, while the Maxim of Quantity specifies that speakers must be as informative as necessary – no more and no less. The Maxim of Relation specifies that speakers must only impart information that is relevant, and the Maxim of Manner specifies that information must be communicated in a clear, brief, and organized way.

When a speaker openly violates a given maxim, the assumption that he is ultimately adhering to the Cooperative Principle leads listeners to assume there must be

some recoverable cooperative purpose behind his violation. This assumption leads listeners to compute implicatures – that is, to infer some additional meaning to the speaker's utterance beyond its literal meaning. For instance, consider the dialogue below (from Craig Raine, 2009):

(1) A: [I want you to] spend the rest of my life with me.

B: The arrangements for that shipment will require detailed advance planning.

A: Is there someone there?

B: That is the correct state of affairs.

The obscure and verbose nature of Speaker B's first utterance clearly violates the maxim of Manner (and perhaps also the maxim of Relevance). However, the assumption that Speaker B is ultimately being cooperative leads Speaker A to infer some additional implicit meaning to Speaker B's utterance – namely, that Speaker A is being secretive because someone else is listening in on their conversation.

Scalar implicatures are specific inferences that arise when a speaker utters a statement as in (2), below, and listeners infer that a stronger alternative utterance in (3) is false – i.e., that not *all* of the buildings were destroyed in the fire.

(2) Some of the buildings were destroyed in the fire.

(3) All of the buildings were destroyed in the fire.

The *not all* meaning derived from (2) is implied but not entailed by this utterance, as evidenced by the fact that it can be canceled without generating a contradiction, as in (4):

(4) Some of the buildings were destroyed in the fire; in fact, all of them were.

By traditional Gricean accounts, this inference involves reasoning about the role of a speaker's knowledge and intentions in uttering a weaker, less informative statement over a stronger, more informative one. Words like *some* and *all* are thought to form scales of increasing strength or informativeness known as *Horn* scales (Horn, 1972), and speakers are generally expected to use the most informative term that they know to be true (given the maxims of Quality and Quantity). Thus, when a speaker utters a statement with a weaker, less informative term on a given Horn scale, as in (2), above, listeners infer that the speaker believes the statement in (3) is false.

The computations involved in this inference can therefore be described in three main steps:

I. Compute the literal meaning of a sentence S

II. Generate a set of relevant scalar alternatives to S

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

On Gricean accounts, this third step – i.e., negating stronger alternatives – can be broken down into two components:

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., $\neg B(p)$.

So, in example (2), we assume that the speaker does not believe the alternative statement in (3) to be true. 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., $B(\neg p)$.

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, and thus believes it to be false (Sauerland, 2004).

The role of epistemic reasoning in scalar implicature computation has been a matter of significant debate, however. For instance, according to Levinson (2000), scalar implicatures are computed by default (i.e., without going through all of the steps described above) and are only later revised or canceled based on contextual or pragmatic considerations. And according to the grammatical view of scalar implicature, these inferences do not involve any reasoning about speaker knowledge or intentionality at all, but instead are computed via a grammatical operator that is triggered by the presence of a lexical scale (Chierchia, Fox, Spector, 2012; Fox, 2007). On this account, scalar implicature involves Steps I-III, above, but the strengthening in Step III occurs via a grammatical operator instead of via epistemic reasoning.

Epistemic reasoning and Scalar implicature in children

Acquisition offers a promising window into understanding the nature of implicature. By exploring how children's epistemic reasoning abilities develop alongside their ability to compute inferences, researchers may be able to determine the precise role this reasoning plays in implicature computation more generally. Following this line of reasoning, a number of studies have investigated children's abilities to compute scalar implicatures and have found that children exhibit striking failures with these inferences until relatively late in development. For instance, studies by Smith (1980), Noveck

(2001), Papafragou and Musolino (2003), and others have shown that children are more willing than adults to accept a statement with the scalar term *some* (e.g., "Some of the horses jumped over the fence") to describe a scene or picture in which *all* is true (e.g., a scene in which all of the horses jumped over the fence) (see also Barner, Chow, & Yang, 2009; Huang & Snedeker, 2009b; Hurewitz, Papafragou, Gleitman, & Gelman, 2006; Musolino, 2004). Similarly, a study by Chierchia et al. (2001) has shown that children between 3 and 6 years of age are significantly more willing than adults to accept utterances containing *or* (e.g., "Each boy has a bike or a skateboard") in contexts where stronger statements containing *and* are true (e.g., each boy has both items).

Children's difficulty with scalar implicature was initially taken as evidence that implicature depends critically on pragmatic competence, since children were thought to lack epistemic reasoning abilities early in development (e.g., false belief; Wimmer & Perner, 1983; Wellman et al., 2001). However, increasing evidence suggests that children have many of the epistemic reasoning abilities required by scalar implicature before their failures with scalar implicature subside. For instance, as early as 9 months of age, children differentiate between a person who deliberately refuses to cooperate and one who is simply unable to, getting frustrated and impatient when an experimenter purposefully withholds a toy from them but not when the experimenter accidentally drops the toy or is unable to reach it (Behne, Carpenter, Call, and Tomasello, 2005). By 12-months of age, children differentiate between knowledgeable and unknowledgeable interlocutors, pointing more to the location of an object when the adult searching for this object is ignorant of its location than when the adult knows exactly where the object is to be found (Liszkowski, et al., 2008). And by 14-18 months, children are much more likely

to imitate intentional actions than unintentional ones (Carpenter, Akhtar and Tomasello, 1998). There is also increasing evidence that children exhibit some basic understanding of false belief before the age of 2 (e.g., Onishi & Baillargeon, 2005; Surian et al., 2007; Southgate et al., 2007).

If children can reason about other people's knowledge, intentions, and beliefs by the time they are 2-4 years of age, and scalar implicatures involve reasoning about a speaker's knowledge, intentions, and beliefs in uttering a weaker, less informative statement over a stronger one, why do children fail to compute these implicatures until as late as 9 years of age? One possibility is that the epistemic reasoning required for computing scalar implicatures involves some understanding beyond that which is involved in early tests of mental state understanding. For instance, although children understand from an early age that other people can have knowledge and beliefs that differ from their own, it may take them a few more years to understand how speakers' knowledge and beliefs affect their specific linguistic choices (e.g., their utterance informativeness) or, conversely, how their specific linguistic choices reflect their mental states. A second possibility is that children's failures with scalar implicature have nothing to do with their epistemic reasoning abilities. After all, by grammatical accounts, scalar implicature does not involve epistemic reasoning; thus, the acquisition of this inference should be unrelated to the development of Theory of Mind and epistemic reasoning. However, even by Neo-Gricean accounts scalar implicature does not just involve epistemic reasoning – it also involves computing the literal meaning of an utterance, generating relevant alternative utterances via Horn scales, and contrasting these alternative utterances with the original statement. Thus, it is possible that children's

failures with scalar implicature stem from difficulty with one or more of these steps rather than from any difficulty with epistemic reasoning. For instance, some researchers have proposed that children fail to compute scalar implicatures because they lack the necessary processing resources (Reinhart, 2004; Chierchia et al., 2001; Pouscoulous et al., 2007). Others have proposed that children's failures stem more specifically from an inability to access the relevant lexical alternatives on a given Horn scale – either because they do not know which items form Horn scales (e.g., they do not know that *all* is a relevant alternative to *some*) or because their representation of Horn scales is weak in early development (e.g. Barner et al., 2011).

In Chapter 2, we attempt to narrow down the possible sources of children's difficulties with scalar implicature by investigating whether 4- and 5-year-olds can compute "ignorance implicatures" – inferences that arise when a speaker makes an underinformative statement like, "Claire plays soccer or tennis" and listeners infer that the speaker is ignorant or uncertain regarding which sport Claire plays. These inferences involve the same reasoning about speaker knowledge and utterance informativeness as that which, by Gricean accounts, is required for scalar implicature (i.e., Steps I-IIIa above), but differ from scalar implicature with respect to the alternative statements that are involved. Furthermore, unlike scalar implicatures (which could theoretically be computed in absence of any epistemic reasoning; e.g., Chierchia et al., 2012), ignorance implicatures *necessarily* require mental state reasoning, as they involve making an inference about a speaker's lack of knowledge. Ignorance implicatures thus provide an ideal area in which to explore children's epistemic reasoning abilities in connection to implicature computation.

We show that 5-year-olds succeed at ignorance implicature yet fail at scalar implicature, suggesting that children have sufficient epistemic reasoning ability to compute basic Gricean inferences in language before they are able to compute scalar implicatures. On the basis of this finding, we argue that children's difficulties with scalar implicature cannot be epistemic in nature and that these difficulties are best explained, instead, by an inability to access the specific lexical alternatives on a given Horn scale. We also find that 4-year-olds fail to compute both ignorance and scalar implicature, suggesting that the ability to compute basic Gricean inferences in language emerges around 5 years of age.

The findings from Chapter 2 therefore help explain the first paradox that we sought to address in this dissertation – i.e., why children fail at scalar implicature despite their early epistemic reasoning abilities. However, they raise new questions about the role of epistemic reasoning in children's earlier inferences in language. Specifically, the finding that 4-year-olds fail to compute ignorance implicature suggests that, at this age, children are unable to reason about speaker knowledge and utterance informativeness to compute basic inferences in language. We would therefore not expect children to compute any inferences based on Gricean reasoning about speaker knowledge and utterance informativeness before this age. However, a few recent studies have argued that children can compute *ad hoc* implicatures by 3 or 4 years of age (e.g., Stiller et al, 2011). *Ad hoc* implicatures are formally identical to scalar implicatures except that the alternative utterances involved are derived from contextual scales (i.e., scales that arise from the context of a conversation) rather than from Horn scales. By Gricean accounts, they involve the same reasoning about speaker knowledge and intentionality as ignorance

implicature. The finding that children compute *ad hoc* implicatures before ignorance implicatures would therefore be somewhat difficult to explain on Gricean accounts. However, the evidence regarding children's ability to compute *ad hoc* implicatures remains inconclusive, and, more importantly, no previous study has directly compared *ad hoc* and ignorance implicature using comparable methods.

In Chapter 3, we examined 4-year-old children's ability to compute both *ad hoc* scalar implicatures and ignorance implicatures using the same methods, to determine whether one precedes the other in development. We show that most 4-year-olds successfully compute *ad hoc* implicatures despite failing to compute ignorance implicatures, which indicates that *ad hoc* implicatures emerge before ignorance implicatures in development. We discuss several possible explanations for this finding, including the possibility that children compute *ad hoc* implicatures without actually engaging in epistemic reasoning.

Epistemic reasoning and Scalar implicature in ASD

Although most of the research on scalar implicature has focused on typically-developing adults and children, researchers have recently turned to Autism Spectrum Disorders (ASD) to explore the role of epistemic reasoning in implicature computation. ASD refers to a cluster of neuro-developmental disorders characterized primarily by deficits in social communication and interaction as well as repetitive behaviors and interests (APA DSM-V, 2013). Linguistic abilities vary widely across the autism spectrum, with some individuals acquiring little to no speech and others exhibiting relatively unimpaired core language ability (e.g., Kjelgaard & Tager-Flusberg, 2001;

Minshew et al., 1995). However, pragmatic deficits appear to be universal in ASD. Even high-functioning individuals with autism (i.e., those with normal cognitive abilities and relatively unimpaired core language abilities) exhibit difficulties with pragmatic aspects of language such as conversational turn-taking, prosody, irony, humor, and metaphor (see Minshew et al., 1995; Baron-Cohen, 1988; Eales, 1993; Happé, 1993; Ozonoff & Miller, 1996; Shriberg et al., 2001; Tager-Flusberg, Paul, & Lord, 2005). For instance, Ozonoff & Miller (1996) presented 17 high-functioning adults with ASD and 17 age- and IQ-matched controls with the first few lines of a joke story and the first few lines of a regular, non-humorous story and asked them to pick one of four possible endings to these stories. They found that the ASD group had significantly more trouble completing the joke endings than the regular short story endings relative to controls, suggesting that even high-functioning individuals with ASD exhibit deficits in the comprehension of humor. Shriberg et al. (2001) examined speech samples from 30 high-functioning children and adults with ASD and Aspergers and 53 typically-developing controls and found significant differences between these groups in phrasing, stress, and nasal resonance. For example, although the ASD group used grammatical and lexical stress appropriately, they exhibited significant deficits with pragmatic and emphatic stress at the utterance level. Finally, Minshew et al., (1995) tested 62 high-functioning adolescents and adults and 50 typically-developing controls on a battery of linguistic tests and found that although the ASD group showed evidence of a “normal lexicon, normal fund of semantic knowledge, and normal grasp of the basic rules of grammar and phonetics” relative to controls, they exhibited significant impairments with “interpretive language abilities” such as making

inferences in language, understanding metaphorical expressions, and interpreting ambiguity (259).

The pragmatic deficits exhibited by individuals with ASD in language have often been linked to an impaired Theory of Mind. For instance, Baron-Cohen et al. (1985) found that high-functioning children with ASD between the ages of 6 and 16 fail at a classic test of false-belief understanding known as the Sally-Anne test, while children with Down Syndrome and typically-developing controls succeed. In this test, subjects watch as a character named Sally places an object in one location, leaves the scene, and then a second character named Anne moves the object to a new location. When asked where Sally will look for the object when she comes back, 80% of the children with ASD in Baron-Cohen et al.'s study pointed to the new location, where the object actually was, despite the fact that Sally never saw the object being moved – suggesting that they failed to appreciate that Sally's knowledge and beliefs differed from their own. Happé (1994) and Baron-Cohen et al. (1997) also found that high-functioning adolescents and adults who pass Sally-Anne type false-belief tasks still perform significantly worse than typically-developing controls on more advanced Theory of Mind tests like the "Strange Stories Test" and the "Reading the Mind in the Eyes" test, which require subjects to reason about the intentions behind various nonliteral uses of language (e.g., white lies, sarcasm, etc.) and to infer a person's mental state by looking at their eyes, respectively. Finally, Happé (1993) found a correlation between individuals with ASD's performance on false-belief and deception tasks and their understanding of non-literal uses of language such as similes, metaphors, and irony, suggesting that impairments with pragmatic aspects of language are associated with Theory of Mind deficits in ASD.

Given these difficulties with pragmatic aspects of language and deficits with Theory of Mind reasoning, one might expect individuals with ASD to fail at pragmatic inferences in language that are thought to require epistemic reasoning, such as scalar implicature. However, two recent studies by Pijnacker et al. (2009) and Chevallier et al., (2010) have argued that high-functioning adolescents and adults with ASD compute scalar implicatures to the same degree as typically-developing controls. For instance, Pijnacker et al. (2009) presented high functioning adults with ASD with sentences like, "Some sparrows are birds", and "Zebras have black or white stripes" and asked them to judge whether these sentences 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 perform differently from typically-developing adult 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 typically-developing controls with under-informative disjunctive 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.

There are at least two possible explanations for this unexpected finding. One is that high-functioning individuals with ASD are capable of engaging in the kind of

epistemic reasoning that is required for scalar implicature, despite their other socio-pragmatic deficits. Specifically, they may be capable of reasoning about the role of knowledge and intentionality in a speaker's decision to utter a weaker, less informative statement over a stronger, more informative one. In this case, we might expect high-functioning individuals with ASD to succeed at other types of pragmatic inference that involve the same kind of reasoning about speaker knowledge and intentionality.

A second possibility is that high-functioning individuals with ASD do not actually engage in any reasoning about speaker knowledge and intentionality when computing scalar implicatures – either because they employ other linguistic or cognitive resources to compensate for their pragmatic deficits (e.g., Pijnacker et al., 2009), or because scalar implicatures do not actually require epistemic reasoning (e.g., Chierchia, et al., 2012). In this case, we might expect high-functioning individuals with ASD to succeed at scalar implicatures yet fail at other pragmatic inferences that require epistemic reasoning.

In Chapter 4, we attempt to tease apart these two possibilities. First, we test whether high-functioning adolescents with ASD are capable of computing ignorance implicatures, since these inferences necessarily require the kind of epistemic reasoning that, by Gricean accounts, is involved in scalar implicature. We use the same methods as those employed in Chapter 2 to test ignorance and scalar implicature computation, and we show that our subjects with ASD succeed at both types of inferences. These results suggest that high-functioning adolescents with ASD are capable of reasoning about the relation between speaker knowledge and utterance informativeness to make inferences in language, despite their other pragmatic deficits.

The fact that our subjects successfully computed ignorance implicatures therefore provides evidence that high-functioning individuals with ASD can engage in the kind of reasoning that, by Gricean accounts, is necessary to compute scalar implicatures; however, it does not address whether individuals with ASD actually engage in such reasoning when computing scalar implicatures. Thus, in Chapter 4, we also examine the role of epistemic reasoning in scalar implicature more directly by testing whether individuals with ASD compute scalar implicatures when given evidence that a speaker has full knowledge about the truth of a stronger alternative utterance (i.e., step IIIb above) but not when given evidence that a speaker only has partial knowledge. We show that high-functioning adolescents with ASD compute scalar implicatures in both contexts, suggesting that they fail to consider the epistemic states of specific speakers when computing scalar implicatures. We discuss various possible explanations for this finding, including the possibility that although individuals with ASD are capable of reasoning about speakers' mental states, they do not always do so when computing scalar implicatures.

Finally, although the results from the first experiment in Chapter 4 indicate that adolescents with ASD are able to compute ignorance implicatures, the results from Chapter 2 show that typically-developing children generally succeed at these inferences much earlier – around 5 years of age. Thus, in Chapter 4, we also tested younger subjects with ASD (between 5 and 10 years of age) on the task from Chapter 1 to determine whether individuals with ASD exhibit a developmental delay in computing implicatures. We show that this younger group failed to compute both scalar and ignorance implicatures despite succeeding at control trials on both tasks, suggesting that the ability

to make Gricean inferences in language is somewhat delayed in ASD relative to typically-developing controls.

Taken together, the results from the studies in Chapters 2 through 4 paint a clearer picture of the development of epistemic reasoning abilities and implicature computation in typically-developing children and in individuals with ASD, and thus help answer the two seeming paradoxes about scalar implicature computation in these two groups that motivated this dissertation. However, our findings also raise new questions about the role of epistemic reasoning in children's earlier inferences, such as *ad hoc* implicature, and about the precise role of epistemic reasoning in scalar implicature computation in ASD. In Chapter 5, we discuss the implications for these findings and suggest directions for future research.

REFERENCES

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

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

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

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

Baron−Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Journal of Child Psychology and Psychiatry*, 38, 813-822.

Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling versus unable: infants' understanding of intentional action. *Developmental Psychology,* 41, 328-337.

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- to 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development,* 21, 315-330.

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*. Cascadilla Press. Somerville. 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.

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

De Villiers, J. (2007). The interface of language and theory of mind. *Lingua*, 117, 1858-1878.

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.

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.

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.

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

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

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

Hurewitz, F., Papafragou, A., Gleitman, L., & Gelman, R. (2006). Asymmetries in the acquisition of numbers and quantifiers. *Language Learning and Development*, 2, 77–96.

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and cognitive processes*, 16, 287-308.

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.

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

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

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

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

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255-258.

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. & 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.

Pouscoulous, N., Noveck, I., Politzer, G., & Bastide, A. (2007). A developmental investigation of processing costs in implicature production. *Language Acquisition*, 14, 347-375.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and brain sciences*, 1, 515-526.

Raine, C. (2009). Love affair with secondaries. *New Yorker*, 76-79.

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

Russell, P. A., Hosie, J. A., Gray, C. D., Scott, C., Hunter, N., Banks, J. S., & Macaulay, M. C. (1998). The development of theory of mind in deaf children. *Journal of Child Psychology and Psychiatry*, 39, 903-910.

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.

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

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18, 587-592.

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.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, 18, 580-586.

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 (Fourth ed., Vol. 1, pp. 335-364). New York: Wiley.

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.

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

CHAPTER 2:

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

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.

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*. 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 2.1, below. Accounts differ only with respect to the mechanism by which the alternatives are negated, as described in Step IV.

Table 2.1: 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., ¬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).

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 *some*, 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* 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.

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.[1] 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

[1] The pragmatic accounts described here might be restated within the grammatical theory. Under such a restatement, pragmatic differences between children and adults (e.g.

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).

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 2.1). 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 2.1 (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*.

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., $\neg B(P)$) nor is it the case that s believes not p (i.e., $\neg B(\neg P)$).

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 2.1; 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.

EXPERIMENT

Methods

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.

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.

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 2.2 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 2.2: 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.[2]	Blindfolded doll
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

[2] 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 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").

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 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 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.

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 2.3 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 2.3: 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.

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.[3] 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 2.1 and 2.2 display performance on the critical and control trials in both tasks by age.

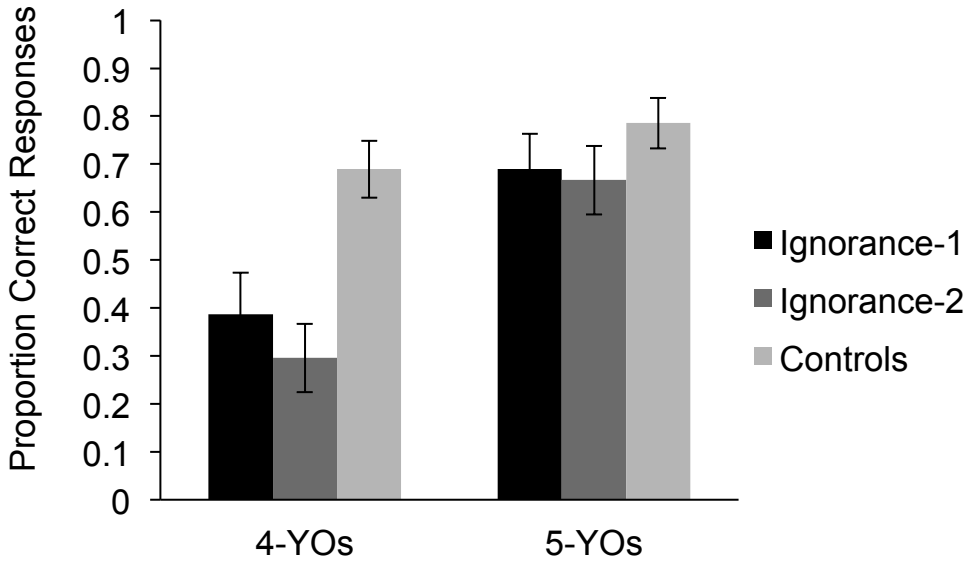

Figure 2.1: Proportion of correct responses on critical and control trials in the Ignorance Implicature task. Error bars represent standard error of the mean.

[3] 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).

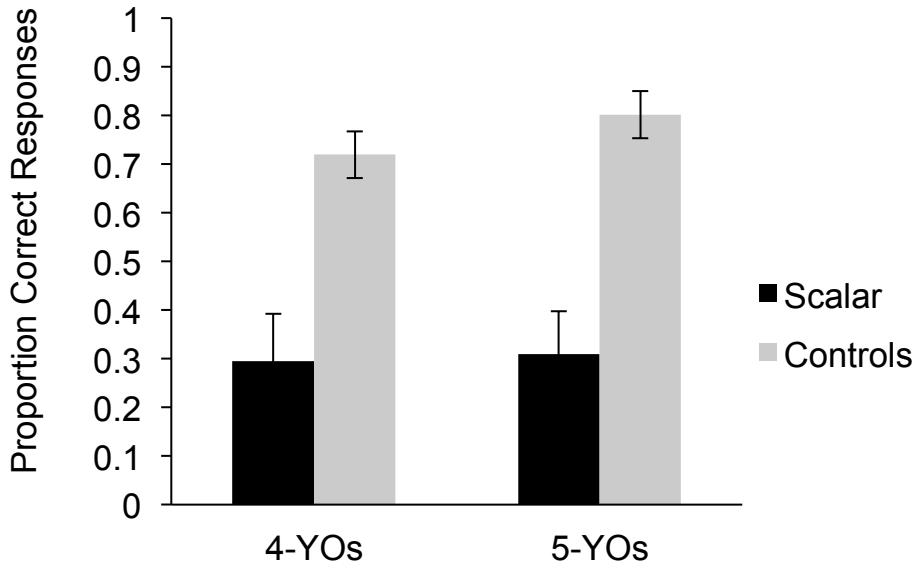

Figure 2.2: Proportion of correct responses on critical and control trials in the Scalar Implicature task. Error bars represent standard error of the mean.

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.

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.[4] 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 basis of the weaker inference that precedes the epistemic step (i.e., $\neg B(p)$). 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

[4] 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.

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)[5]. 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., "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

---

[5] 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.")

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).[6] 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.[7] 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

---

[6] The proposal of Franke (2011) shares with the Neo-Gricean approach the view that SI are pragmatic in nature, but differs in that SIs are computed by different mechanisms than those relevant for the computation of ignorance inferences, as in the grammatical theory. Though Franke does not focus on the computation of ignorance inferences, we assume that he could develop his theory so that the mechanism that generates SIs (the base-level interpretation game) is based on a different set of alternatives than the one relevant for computing ignorance inferences (the epistemic interpretation game). If that is the case, and if children's difficulty with SIs is to be explained in terms of a difficulty with alternatives, the grammatical theory and Franke's would make the same predictions – or so it seems to us.

[7] 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).

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 Hirschberg, 1985).[8] 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.

---

[8] 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."

(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 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 in 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.

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.

Acknowledgements

Many thanks to Dave Barner, Andy Kehler, Jess Sullivan, Ellie Chestnut, Katie Kimura, Annalice Heinz, Troy Campbell, Laura Taouk, 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.

Chapter 2, in full, is a reprint of the material currently under review at the *Journal of Semantics*, and may appear as Hochstein, L., Bale, A., Fox, D., & Barner, D. Ignorance and Inference: Do problems with Gricean epistemic reasoning explain children's difficulties with scalar implicature? The dissertation author was the primary investigator and author of this paper. Permission to reprint was granted by the *Journal of Semantics.*

REFERENCES

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

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.

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

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

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

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

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

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

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

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. (2009). Hurford's constraint and the theory of scalar implicatures. *Presuppositions and implicatures*, 60, 47-62.

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.

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

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

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.

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

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

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

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

Goro, T., Minai, U., & Crain, S. (2005). Two disjunctions for the price of only one. *Proceedings of the 29th Boston University Conference on Language Development, Somerville.* MA: Cascadilla Press.

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

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

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

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

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

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

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

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

Hurewitz, F., Papafragou, A., Gleitman, L., & Gelman, R. (2006). Asymmetries in the acquisition of numbers and quantifiers. *Language Learning and Development*, 2, 77–96.

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

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

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

Marty, P., Chemla, E., & Spector, B. (in press). Interpreting numerals and scalar items under memory load. *Lingua*.

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

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

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

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

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

Pouscoulous, N., Noveck, I., Politzer, G, & Bastide, A. (2007). A developmental investigation of processing costs in implicature production. *Language Acquisition*, 14, 347-375.

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

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

Russell, B. (2011). Topics in the Computation of Scalar Implicatures. PhD dissertation, Brown University. Posted on Semantics Archive: http://semanticsarchive.net/Archive/WY1YTRhM/russell2012probabilistic.pdf

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

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.

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

Stiller, A., Goodman, N., & Frank, M. (under review). Ad-hoc scalar implicatures in adults and children.

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

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

Wynn, K. (1992). Children’s acquisition of number words and the counting system. Cognitive Psychology, 24, 220–251.

CHAPTER 3:

Epistemic reasoning in children's early linguistic inferences: the case of *ad hoc* implicature

Abstract

Increasing evidence suggests that children have relatively sophisticated epistemic reasoning abilities from an early age, however it remains unclear when they actually begin to apply this understanding to make inferences in language. In this paper, we examine children's ability to compute two inferences, which, by most accounts, involve the same epistemic reasoning abilities, but which have been claimed to be acquired at different ages: ignorance implicature and *ad hoc* scalar implicature. Unlike previous studies, which have tested the two inferences separately, using very different methods, we examined 4-year-old children's ability to compute both *ad hoc* scalar implicatures and ignorance implicatures using the same methods, to determine whether one really does precede the other in development. We show that most 4-year-olds successfully compute *ad hoc* implicatures despite failing to compute ignorance implicatures, which suggests that *ad hoc* implicatures emerge before ignorance implicatures. We discuss several possible explanations for this finding, including the possibility that children compute *ad hoc* implicatures without actually engaging in epistemic reasoning.

INTRODUCTION

To become competent language users, children must not only acquire the literal meanings of words but also make inferences about speakers' intended meaning – a process that is generally thought to require epistemic reasoning about the speaker's knowledge and intentions (e.g., Grice, 1975). Although there is evidence that children can reason about other people's mental states from an early age (e.g., O'Neill, 1996; Behne et al., 2005; Carpenter et al., 1998; Baillargeon et al., 2010), it remains unclear exactly when children begin to apply this understanding to enrich the literal meanings of utterances pragmatically. Characterizing the development of linguistic pragmatic competence is crucial for determining the role of pragmatic reasoning abilities in language acquisition, as some theorists argue that such abilities actually drive language acquisition (e.g., Bloom, 2000; Tomasello, 2003; Baldwin and Meyer, 2007) while others argue that they are secondary and do not emerge until relatively late in development (e.g., Noveck, 2001; Wexler, 2003). In this paper, we explore children's ability to compute *ad hoc* scalar implicatures, pragmatic inferences that have been claimed to be acquired relatively early in development, and we ask whether these implicatures plausibly involve epistemic reasoning.

Over the first two years of life, children develop a basic understanding of other people as intentional beings who can share or lack knowledge. For instance, as early as 9 months of age, children differentiate between a person who deliberately refuses to cooperate and one who is simply unable to (Behne et al., 2005), and by 12 months of age, children point more to the location of an object when the adult searching for this object is ignorant of its location than when the adult knows exactly where the object is to be found

(Liszkowski et al., 2008). By 14 to 18 months, children are much more likely to imitate intentional actions than unintentional ones (Carpenter, et al., 1998). Recently, a number of researchers have even argued that children have some understanding of false belief before the age of 2 (e.g., Onishi & Baillargeon, 2005; Surian et al., 2007; Southgate et al., 2007). For instance, Onishi and Baillargeon (2005) found that 15-month-olds look longer at scenes in which an agent searches for an object in the wrong place despite knowing the object's correct location than at equivalent scenes in which the agent did not know the object's correct location. And, conversely, 15-month-olds look longer at scenes in which an agent searches for an object in the correct location despite previously having seen the object in a different location (i.e., despite having a false belief about the location of the object) than at equivalent scenes in which the agent knew the object's correct location. These results suggest that even 15-month-olds expect agents to act in accordance with their beliefs – whether true or false – and are surprised when they do not.

There is thus increasing evidence that children have relatively sophisticated epistemic reasoning abilities from an early age. However, it remains a matter of significant debate when children actually begin to apply this understanding of other people's knowledge, beliefs, and intentions to make inferences in the service of language. For instance, by the age of 2, children can compute inferences like mutual exclusivity (or contrast) which, by some accounts, depend critically on epistemic reasoning – specifically, on the ability to reason about a speaker's intentions (Clark, 1990; Gathercole, 1989; Bloom, 2000; Diesendruck & Markson, 2001; Woodward & Markman, 1998). According to such accounts, when shown two objects – one known and one novel – children assume that an unfamiliar word applies to the novel object based on the

Gricean assumption that a speaker would have used a familiar word if he had wanted to reference a familiar item. The child's reasoning would therefore be something along the lines of, "If the experimenter had wanted me to pick up the cup, she would have asked me to show her the cup. Since she asked me for the dax instead, I infer that she must want me to give her the other object" (e.g., Diesendruck & Markson, 2001). However, the role of epistemic reasoning in mutual exclusivity has been a matter of significant debate, with some researchers arguing that this inference is based simply on the assumption that objects only have one label, and therefore does not specifically involve reasoning about the epistemic states of speakers (e.g., Markman & Wachtel, 1988; de Marchena et al, 2011; Preissler & Carey, 2005).

To further explore the role of epistemic reasoning in children's early inferences in language, researchers have investigated children's ability to compute scalar implicature, an inference that, by some accounts, is formally similar to Mutual Exclusivity (e.g., Barner & Bachrach, 2010). Scalar implicatures arise when a speaker utters a statement as in (1), and listeners infer that the stronger alternative statement in (2) is false.

(1) Saya bought some of Dimitri's paintings.

(2) Saya bought all of Dimitri's paintings.

By standard Gricean accounts (e.g., Grice, 1975; Horn, 1972; Gazdar, 1979; Geurts, 2009; Russell, 2011), this inference is based, first, on the assumption that words like *some* and *all* form scales of increasing strength or informativeness (known as *Horn scales*; Horn, 1972) and, second, on the assumption that the speaker is being truthful and cooperative and is providing an appropriate amount of information given what he knows to be true (Grice, 1975). Thus, if the speaker knew that Saya bought all of Dimitri's

paintings, he would have simply said so. Because he chose not to utter this stronger, more informative statement, listeners infer that he does not know or believe it to be true. Just as with mutual exclusivity, however, the role of epistemic reasoning in the computation of scalar implicature has come under debate, with some theorists arguing that this inference is based on a grammatical process that does not involve any reasoning about a speaker's knowledge and intentions (e.g., Chierchia, Fox, & Spector, 2012; Chierchia 2004, 2006; Fox 2007, Fox & Hackl 2006; Landman 1998, among others).

A number of studies have shown that, surprisingly, children up to 9 years of age fail to compute scalar implicatures (e.g., Smith, 1980; Noveck, 2001; Papafragou and Musolino, 2003). Recent work has explored whether these late failures are due to difficulty with the epistemic reasoning involved in these inferences (e.g., Katsos & Bishop, 2011), or to other factors like general processing limitations (e.g., Reinhart, 2004; Chierchia et al., 2001; Pouscoulous et al., 2007) or an inability to access the specific alternatives in a given Horn scale (e.g., Barner & Bachrach, 2010; Barner et al., 2011; Chierchia et al., 2001, among others). In relation to this, Hochstein et al. (under review) have argued that children's difficulties are not likely epistemic in nature. In support of this, they showed that although 5-year-olds fail to compute scalar implicatures, they successfully compute ignorance implicatures – inferences which, on Gricean accounts, are required as part of the process of computing scalar implicatures, and which, on all accounts, involve reasoning about the epistemic states of the speaker. Ignorance implicatures arise when, for instance, a speaker utters a disjunctive statement like the one in (3), and listeners infer ignorance or uncertainty on the part of the speaker regarding exactly where each of her friends went for vacation.

(3) Each of my friends went to Peru or Brazil for their vacation.

(4) Each of my friends went to Peru for their vacation.

(5) Each of my friends went to Brazil for their vacation.

(6) Each of my friends went to Peru and Brazil for the vacation.

This inference is based on the assumption that if the speaker in (3) knew that each friend went to Peru, she should have simply uttered the stronger statement in (4). Similarly, if the speaker knew that each friend went to Brazil, she should have uttered the stronger statement in (5). Finally, if she knew that each friend went to both places, she should have said so, by uttering the stronger statement in (6). Since the speaker did not utter either of these stronger and more informative alternative utterances, listeners infer that she does not know any of them to be true. This inference, the ignorance implicature, is represented in (7).

(7) $\neg B(p) \wedge \neg B(q) \wedge [\neg B(p \wedge q)]$

(8) $B(\neg p) \wedge B(\neg q) \wedge [B\neg (p \wedge q)]$

By most Gricean accounts, scalar implicatures differ from ignorance implicatures due to an additional inference, called the "epistemic step" (e.g., Sauerland, 2004), whereby the listener infers that the speaker is knowledgeable, and thus that she believes the stronger statements to be false (as in 8). Thus, in (3), the listener infers that each friend went to either Peru or Brazil, but not both, and that not everyone went to the same place.

The fact that children can compute ignorance implicatures before scalar implicatures strongly suggests that their failures with scalar implicature cannot be

attributed to a difficulty with this type of epistemic reasoning.[9] However, while Hochstein et al. found that 5-year-olds could compute ignorance implicatures, they also found that 4-year-olds generally failed to compute these inferences. This finding is interesting, given the substantial evidence (discussed above) that children are able to reason in other relatively sophisticated ways about other people's knowledge, intentions, and beliefs by the time they are 2 years of age. The fact that children do not appear to compute ignorance implicatures until much later suggests that this inference involves some understanding beyond that which is involved in evaluating another person's mental state. Computing ignorance implicatures also involves understanding how a speaker's mental state affects utterance informativeness. Thus, although children begin to reason about intentional states very early in development, it may take them much longer to understand the connection between speaker knowledge and utterance informativeness and to make inferences on the basis of this understanding.

On the Gricean hypothesis that scalar implicature depends upon the ability to reason about speaker knowledge (and specifically, to compute ignorance implicatures), we would not expect children to compute scalar implicatures before the age at which they first compute ignorance implicatures. However, recent studies have argued that children

---

[9] Also, as Hochstein et al. argue, children's difficulty with scalar implicature cannot be attributed to a failure to take the "epistemic step." In their study, the knowledge states of speakers were provided contextually and did not need to be inferred; the evidence required for taking the epistemic step was thus made readily available to subjects. Furthermore, if children's difficulty with scalar implicature stemmed solely from a failure to take the epistemic step, we would still expect them to compute ignorance implicatures for underinformative statements. However, in Hochstein et al.'s study, children attributed underinformative statements involving scalar implicature to smart, knowledgeable speakers, suggesting that they did not compute either scalar or ignorance implicatures for these statements.

can compute *ad hoc* scalar implicatures, inferences that are based on contextual scales, by 3 or 4 years of age (e.g., Stiller et al, 2011; Papafragou & Tantalou, 2004). By most accounts, *ad hoc* implicatures are formally identical to scalar implicatures except that the alternative utterances involved are derived from contextual scales (i.e., scales that arise from the context of a conversation) rather than from Horn scales (e.g., Hirschberg, 1985). For instance, consider the dialogue in (9):

(9) Speaker A: Are Monica, Jess, and Abby going to the party?

Speaker B: Monica is going to the party.

In (9), we infer that Jess and Abby are not going to the party. By standard Gricean accounts, this inference is based on the assumption that if Speaker A knew that Jess and Abby were going to the party, she would have simply said so; it therefore involves the same reasoning about speaker knowledge and intentionality as ignorance implicature. Thus, evidence that children can compute *ad hoc* implicatures before ignorance implicatures would be somewhat difficult to explain on Gricean accounts.

The evidence that children can compute *ad hoc* implicatures by 3 or 4 years of age remains inconclusive, however. For example, in one study sometimes taken to demonstrate an early ability to compute *ad hoc* implicatures, Barner et al. (2011) found that 4-year-old children can access alternatives based on *ad hoc* scales more readily than they can access members of Horn scales like *some* and *all*. In their study, children were shown a picture of a cat, dog, and cow sleeping, and asked either “Are only some of the animals sleeping?” or “Are only the cat and the dog sleeping?”. Children said “yes” to the former question but “no” to the latter, whereas adults said “no” to both, because in each case the statement was logically false. Critically, though, no implicature computation was

involved in this task, since the word "only" entails the falsity of stronger statements, as demonstrated by the fact that statements involving "only" are not defeasible (e.g., it is not possible to say, "Only some of the animals are sleeping; in fact, all of them are").[10] This study therefore provided evidence that children can access alternatives based on *ad hoc* scales before they can access Horn scale alternatives, but it did not address whether or not children could actually compute implicatures based on *ad hoc* scales.

Stiller et al. (2011) set out to test *ad hoc* implicature computation more directly using a forced choice paradigm. They presented 3- and 4-year old children with sets of three pictures that differed in terms of the number of features they exhibited. For instance, children were shown three smiley faces: one with no accessories, one with only glasses, and one with glasses and a top hat – thus forming the following contextual scale: <glasses, a hat, a hat and glasses>. The experimenter then told children, "My friend has glasses. Can you show me my friend?" The authors argued that, in this context, "My friend has a top hat" would have been a stronger and more informative description of the face with glasses and a top hat. Thus, the utterance "my friend has glasses" implies that the friend does not have a top hat. And, indeed, children as young as 3 years of age in this study reliably chose the face with only glasses in this context. However, utterances like, "My friend has glasses. Can you show me my friend" presuppose that *only* one of the three possible referents is the correct one. Thus, children could have chosen the face with only glasses not because they ruled out the face with glasses and a top hat as a potential

[10] Furthermore, questions are generally not believed to generate implicatures, a fact which was also borne out by Barner et al.'s study in contexts where the word "only" was omitted from questions. When asked, "Are some of the animals sleeping," both children and adults said "yes" even if all of the animals were sleeping.

referent (via implicature), but because they had a slight preference for the face with glasses when forced to choose between the two (e.g., because it most closely matches the description they were given). We know that when presented with two alternative descriptions of a scene and forced to choose one, children prefer the stronger, more informative one, yet they are willing to accept the weaker, less informative description of the same scene when it is presented in isolation (Chierchia et al., 2001). Thus, although the forced choice task used by Stiller et al. demonstrates that children are sensitive in some way to utterance informativeness, it does not show whether children actually consider a statement like "My friend has glasses" to be an infelicitous description of the man with glasses and a hat (see also Barner & Bachrach, 2010, and Barner et al., 2011).

Finally, in a study by Papafragou and Tantalou (2004), 4- to 6-year-old children were presented with a scene in which a character was instructed to complete a task (e.g., to wrap a toy parrot and a doll). When the character was asked whether it had completed the task (e.g., "Did you wrap the gifts?"), it responded by mentioning only one of the required items (e.g., "I wrapped the parrot"). Children were then instructed to reward the character only if it had successfully completed the task. Papafragou and Tantalou found that their subjects refused to reward the character 90% of the time on these trials, which they took as evidence that 4- to 6-year-olds can successfully compute *ad hoc* implicatures (i.e., children interpreted "I wrapped the parrot" as implying the falsity of the stronger alternative utterance, "I wrapped the parrot and doll"). However there are several problems in interpreting this finding. First, relevant to the present study, the mean age of their subjects was 5;3, making it unclear whether 4-year-olds can generally compute these implicatures. Furthermore, in a subsequent study, Sullivan et al. (2011) provided

compelling evidence that children's behavior on Papafragou and Tantalou (2004)'s task may not actually reflect implicature computation. In their study, Sullivan et al. replicated Papafragou and Tantalou's methods and found that in addition to refusing rewards for characters who only mentioned a subset of the required items, children also refused rewards for characters who mentioned *more* items than were required. For instance, if a character was instructed to wash a hat, was asked, "Did you wash the hat?" and then answered, "I washed the hat and the shirt," children refused to reward the character, despite the fact that the character had clearly met the requirements of the task. Thus children not only rejected statements that implied the falsity of stronger alternative utterances, they also rejected statements that entailed the truth of weaker alternative utterances. These findings suggest that children simply rejected the characters' utterances whenever they contrasted in some way with the original request – and not because they computed implicatures (for discussion, see also Bale & Barner, 2013).

It remains unclear, then, exactly when children can successfully compute *ad hoc* implicatures. And although results from previous studies seem to suggest that *ad hoc* implicatures emerge before ignorance implicatures, no previous study has directly compared these two forms of inference using comparable methods. To address this, in the present study we examined 4-year-old children's ability to compute both *ad hoc* scalar implicatures and ignorance implicatures in almost identical situations, to determine whether the ability to compute *ad hoc* implicatures precedes the ability to make Gricean epistemic inferences. To this end, we tested 4-year-old children on two matched tasks: one assessing *ad hoc* implicature computation and one assessing ignorance implicature computation. We found that 4-year-olds generally fail at ignorance implicature yet

succeed at *ad hoc* implicature, raising the possibility that *ad hoc* implicatures, at least initially, do not depend on Gricean epistemic reasoning.

EXPERIMENT

Method

Participants

We tested 50 monolingual English-speaking 4-year-olds (25 females, *M* = 4;4, range = 4;0-4;11), recruited either by phone or through daycares in San Diego, CA and Comox, B.C. Two additional children were excluded; one due to experimenter error and the other due to failure to complete the task.

Procedure & Stimuli

Subjects were seated at a small table next to the experimenter and were administered one of two tasks: an *Ad hoc Implicature* task and an *Ignorance Implicature* task. In each task, subjects watched a scene on a computer, were presented with a sentence, and were asked to determine which of two characters most likely uttered the given statement.

Ignorance 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 four warm-up trials, there were three or four objects on the lower right-hand corner of the screen (e.g., 3 toy cars) and an animal (e.g., a cow) in the upper left-hand corner. The animal introduced itself, identified the items on the screen, announced its intention to take something, and then took all but one of the items (e.g., "It's me, cow! Look, 3 cars! Look what I'm taking!"). The experimenter then looked at Farmer Brown and Captain Blue and said, "Hey puppets, what did the [cow] take?" while leaning in to "listen" to the puppets whisper their response. On two of the trials, subjects were then presented with a sentence that mentioned more items than were actually taken (e.g., "The cow took 3 cars" when in fact it took 2) and were asked, first, whether this was "the best thing to say" and, second, whether it was Captain Blue or Farmer Brown who uttered this sentence. On the other two trials, subjects were presented with a sentence that accurately described the number of items the animal had taken (e.g., "The cow took 2 cars") and were asked, first, whether this was "the best thing to say", and, second, whether it was Captain Blue or Farmer Brown who uttered it. These warm-up trials were designed to familiarize subjects with attributing sentences to one or the other character. Subjects were given feedback on their performance on these trials.

The 10 test trials were identical to the warm-up trials, except that there were two different objects on the lower right-hand corner of the screen (e.g., a banana and a toy car) and on each trial the animal either took one or both of the items. These test trials consisted of 5 different types, as shown in Table 3.1 below, and these trials were

presented in one of two counter-balanced orders. There were four critical trials and two trials for each control trial-type (four total).

Table 3.1: Test trials in the Ignorance Implicature Task

Condition	Choices	Animal Takes	Someone Says…	Correct Response
True (Control)	Orange / Banana	Orange & Banana	The bear took an orange and a banana.	Seeing doll
False (Control)	Orange / Banana	Orange	The bear took an orange and a banana.	Blindfolded doll
Ignorance (Critical)	Orange / Banana	One or both items	The bear took an orange or a banana.	Blindfolded doll

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* statements were literally true, we expected that subjects would attribute them to the blindfolded doll if they were able to compute ignorance implicatures.

Ad Hoc Implicature Task

This task was identical to the *Ignorance Implicature* task, except that subjects attributed sentences to smart vs. silly speakers rather than to knowledgeable vs. ignorant ones, and the critical trials differed slightly (see Table 3.2 below). 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.

Table 3.2: Test trials in the Ad Hoc Implicature Task

Trial Type	Choices	Animal Takes	Someone Says…	Correct Response
True (Control)	Orange / Banana	Orange & Banana	The bear took an orange and a banana.	Smart puppet
False (Control)	Orange / Banana	Orange	The bear took an orange and a banana.	Silly puppet
Ad Hoc (Critical)	Orange / Banana	Orange & Banana	The bear took an orange.	Silly puppet

The *True* trials were attributable to the Smart Puppet and the *False* trials to the *Silly Puppet.* Though literally true, we expected the *ad hoc* sentences to be attributed to the *Silly Puppet* if children were able to compute *ad hoc* implicatures.

Results

Prior to analysis, we compared performance on *True* and *False* control trials within each task and found no 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. Figure 3.1 displays performance on the critical and control trials in both tasks.

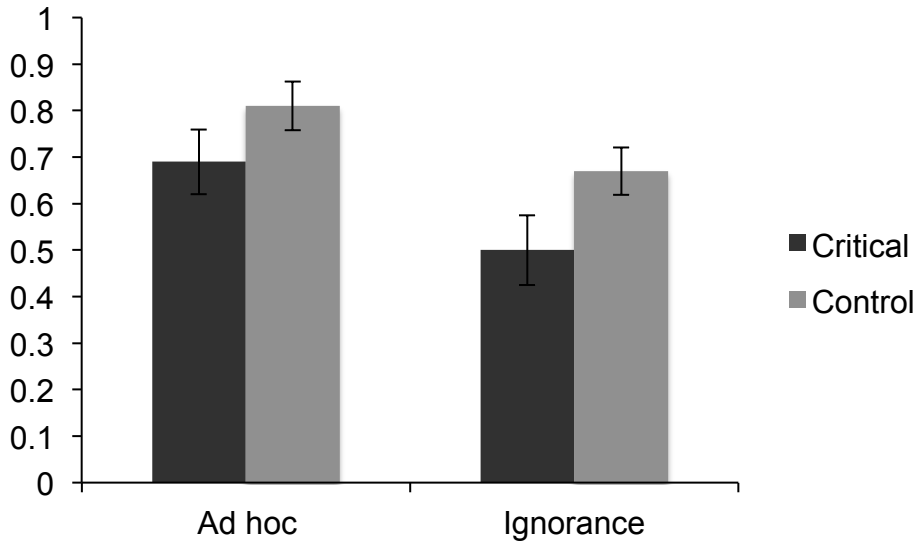

Figure 3.1: Proportion of correct responses on critical vs. control trials in the Ad hoc and Ignorance Implicature tasks. Error bars represent standard error of the mean.

The main question we sought to answer was whether 4-year-olds exhibit evidence of computing *ad hoc* implicatures before ignorance implicatures, or vice versa. To this end, we used a binomial mixed effects regression to predict binary response (correct vs. incorrect) using Task (Ignorance Implicature vs. Ad hoc Implicature) and Trial Type (Critical vs. Control) as within-subjects fixed variables and subject as a random factor. This model found a significant main effect of Task ($z = -2.25$; $p = .024$) and Trial Type ($z = -2.16$, $p = .031$), but no significant interaction ($z = .02$, $p = .99$), indicating that subjects performed better overall on the *Ad hoc Implicature* task than on the *Ignorance Implicature* task, and that they performed better overall on the control trials than on the critical trials in the two tasks. There was no significant difference between critical ($M = .69$, $SE = .070$) and control trials ($M = .81$, $SE = .053$) for the *Ad hoc Implicature* task

(Wilcoxon T = 13.5, N = 11, *p* = .087), and only a marginally significant difference between critical (*M* = .5, *SE* = .075) and control trials (*M* = .67 *SE* = .051) for the *Ignorance Implicature* task (Wilcoxon T = 66, N = 21, *p* = .087).

Finally, we compared subjects' performance on the control trials across the two tasks and found a marginally significant difference in performance (Mann-Whitney U = 413.5, p = .05), suggesting that subjects performed slightly better on control trials in the *Ad hoc Implicature* task than on control trials in the *Ignorance Implicature task*.

Taken together, these results suggest that the 4-year-olds in our study performed better overall on the *Ad hoc Implicature* task than on the *Ignorance Implicature* task.

DISCUSSION

In our experiment, most 4-year-olds successfully computed *ad hoc* implicatures despite failing to compute ignorance implicatures. This finding suggests that *ad hoc* implicatures may emerge before ignorance implicatures, which raises some interesting questions about the role of epistemic reasoning in the computation of *ad hoc* implicature (although, as we discuss further below, there are several possible objections to this interpretation). Specifically, the fact that 4-year-olds fail at ignorance implicature suggests that they do not yet understand the relation between speaker knowledge and utterance informativeness, since ignorance implicature necessarily requires this kind of Gricean reasoning. And if 4-year-olds do not have truly Gricean reasoning abilities yet, then they may not be using such reasoning to compute *ad hoc* implicatures.

If children are not using Gricean epistemic reasoning to compute *ad hoc* implicatures, how are they generating these inferences? One possibility is that *ad hoc*

implicatures are computed grammatically. As mentioned briefly in the Introduction, grammatical accounts of scalar implicature have argued that this inference does not involve any epistemic reasoning. Instead, it is computed via a phonologically null operator whose meaning is similar to that of the word *only* in English (e.g., Chierchia, Fox, & Spector, 2012). It is possible, then, that a similar grammatical process is involved in the computation of *ad hoc* implicatures. However, according to grammatical accounts of scalar implicature, the relevant operator is triggered by the presence of a lexical scale (i.e., <all, some> or <and, or>). Thus, it remains unclear how such an operator could be triggered in *ad hoc* implicatures, which are based on contextually-defined scales rather than lexical scales.

Another possibility is that, like Mutual Exclusivity, *ad hoc* implicature is computed via more general assumptions about language use rather than inferences about a specific speaker's epistemic state. For instance, it has been argued that Mutual Exclusivity does not involve reasoning about a specific speaker's intentions in uttering a novel word but instead involves reasoning more generally about the way words are used and understood by speakers of a given language (e.g., Markman & Wachtel, 1988). Thus, when children hear a novel word like "blick" in the presence of two potential referents – e.g., an apple and a novel object – they assume that "blick" refers to the novel object because of a general expectation that objects are only assigned one label. Since they already know the label for "apple", they infer that "blick" must be the label for the novel object. Similarly, listeners may compute *ad hoc* implicatures based on general expectations about utterance informativeness without actually considering a specific speaker's mental state. For instance, in a dialogue like the one in (4), listeners may

assume that Jess and Abby are not going to the party simply because they assume that speakers are generally maximally informative and thus the statement, "Monica is going to the party" is maximally informative – and not because they are reasoning more specifically that Speaker B would have uttered the stronger statement, "Monica, Jess, and Abby are going to the party" if she had known it to be true.

To test this second possibility more directly, future studies should investigate whether children and adults are sensitive to the knowledge states of speakers when computing *ad hoc* implicatures. As mentioned in the Introduction, 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, if listeners compute *ad hoc* scalar implicatures by reasoning about a specific speaker's knowledge and intentions in choosing a weaker utterance over a stronger alternative one, then they should not compute these inferences when given evidence that the speaker does not know whether the stronger alternative utterance is true. However, if subjects compute *ad hoc* implicatures instead via a general assumption that speakers are maximally informative, then they may compute these inferences regardless of the speaker's epistemic state. Adults' sensitivity to speakers' knowledge states when computing *ad hoc* implicatures could be assessed using a reading-time paradigm like the one in Bergen and Grodner (2012), for instance.

There are a few possible objections to our claim that children compute *ad hoc* implicatures before ignorance implicature, however. First, it is possible that 4-year-olds' success on the *Ad hoc implicature* task in our study does not actually reflect an ability to compute implicatures. For instance, it is possible that subjects did not actually have to

negate a stronger alternative utterance to succeed on this task, but instead attributed underinformative utterances to the silly puppet simply because they noticed a discrepancy between the number of items taken on a given trial and the number of items mentioned (e.g., the bear took a cup and a plate and someone said, “The bear took a cup”). We find this explanation somewhat unlikely, though, given that children did not reliably attribute utterances to the blindfolded doll in the *Ignorance Implicature* task on trials in which there was where a discrepancy between the number of items taken and the number of items mentioned (e.g., when an animal took both items and only one item was mentioned). It is also possible that the *Ignorance Implicature* task was somehow harder than the *Ad hoc Implicature* task. After all, subjects performed worse overall on the *Ignorance Implicature* task (i.e., on both critical and control trials), and there was thus no significant interaction between task (*Ignorance* vs. *Ad hoc*) and trial type (critical vs. control) in our experiment. However, the control trials in the *Ignorance Implicature* and *Ad hoc Implicature* tasks were identical; on the *True* trials an animal took both available items (e.g., the cup and the plate) and children were presented with an utterance that correctly described what had happened (e.g., “The animal took a cup and a plate”) and on the *False* trials an animal took only one of the available items and children were presented with an utterance that was incorrect because it mentioned both items (e.g., “The animal took a cup and a plate”). The only difference between the control trials in these two tasks was the fact that in the *Ad hoc Implicature* task, subjects were asked to attribute these utterances to a smart puppet who always said things that were “just right” or to a silly puppet who always said things that were “silly”, whereas on the *Ignorance Implicature* task they were asked to attribute these utterances to a doll who could see or

one who was blindfolded. In the *Ad hoc Implicature* task, then, children may have simply noticed whether an utterance was "smart" or "silly" and then picked the corresponding puppet. In the *Ignorance Implicature* task, however, subjects had to reason more explicitly about the role of speaker knowledge to conclude that the seeing doll was the only one who had sufficient evidence to make true statements, and the blindfolded doll was the only one in a position to guess incorrectly about what the animal took. Thus, subjects may have performed worse overall on the *Ignorance Implicature* task precisely because this task involved epistemic reasoning, even on the control trials.

In our study, however, we tested different subjects on the *Ignorance Implicature* and *Ad hoc Implicature* tasks. Thus, it is possible that the children we sampled in Experiment 1 happened to be more advanced overall than the children we sampled in Experiment 2. To rule out this possibility, future studies should test 4-year-olds on these two implicature types using a within-subjects design. Future studies should also explore the role of children's developing Theory of Mind (especially understanding of false belief) in their acquisition of *ad hoc* and ignorance implicature computation.

To summarize, in this study we explored children's ability to compute *ad hoc* scalar implicatures and ignorance implicatures by directly comparing these two forms of inference for the first time using comparable methods. We found that 4-year-olds successfully compute *ad hoc* implicatures despite failing to compute ignorance implicatures. Since ignorance implicatures necessarily require reasoning about a speaker's epistemic state, this finding raises the possibility that children compute *ad hoc* implicatures before they are able to engage in epistemic reasoning about speaker knowledge and utterance informativeness. However, this conclusion remains tentative,

given that we tested different subjects on the two tasks and subjects performed worse overall on the *Ignorance Implicature* task, and thus more research is needed to explore the precise role of epistemic reasoning in *ad hoc* implicature.

Acknowledgments

Many thanks to Dave Barner, Katie Kimura, Laura Taouk, 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.

Chapter 3, in full, is currently being prepared for submission for publication. The dissertation author was the primary investigator and author of this material

REFERENCES

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in cognitive sciences*, 14, 110-118.

Baldwin, D. & Meyer, M. (2007). How inherently social is language? In E. Hoff & M. Shatz (Eds.), *Blackwell Handbook of Language Development.* Malden, MA: Blackwell Publishing, 87-106.

Bale, A. & Barner, D. (2013). Grammatical alternatives and pragmatic development. In Anamaria Falaus (Ed.), *Alternatives in Semantics*, Studies in Pragmatics, Language and Cognition, Palgrave.

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.

Behne, T., Carpenter, M., Call, J., & Tomasello, M. (2005). Unwilling versus unable: infants' understanding of intentional action. *Developmental Psychology,* 41, 328-337.

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

Bloom, P. (2000). *How Children Learn the Meanings of Words.* Cambridge, MA: MIT Press.

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- to 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development,* 21, 315-330.

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

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

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*. Cascadilla Press. Somerville. 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.

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

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.

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

Gathercole, V. C. (1989). Contrast: A semantic constraint? *Journal of Child Language*, 16, 685–702.

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

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

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

Hirschberg, J. (1985). *A theory of scalar implicature*. Ph.D. dissertation. University of Pennsylvania.

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

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.

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

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

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. Cognitive Psychology, 20, 121-157.

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

O'Neill, D. K. (1996). Two‐Year‐Old children's sensitivity to a parent's knowledge state when making requests. *Child development*, 67, 659-677.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255-258.

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

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

Pouscoulous, N., Noveck, I., Politzer, G., & Bastide, A. (2007). A developmental investigation of processing costs in implicature production. *Language Acquisition*, 14, 347-375.

Preissler, M. A., & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition*, 97, B13-B23.

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

Russell, B. (2011). Topics in the Computation of Scalar Implicatures. PhD dissertation, Brown University. Posted on Semantics Archive: http://semanticsarchive.net/Archive/WY1YTRhM/russell2012probabilistic.pdf

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

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

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18, 587-592.

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*.

Sullivan, J., Davidson, K., & Barner, D. (2011). Children's conversational implicatures. BUCLD 36, Boston, MA.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, 18, 580-586.

Tomasello, M. (2003). *Constructing a Language: A Usage-Based theory of Language Acquisition.* Cambridge, MA: Harvard University Press.

Wexler, K. (2003). Cues don't Explain Learning: Maximal Trouble in the Determiner System. In T. Gibson & N. Pearlmutter (Eds.), *Proceedings of the 2003 CUNY Sentence Processing Conference.*

Woodward, A. L., & Markman, E. M. (1998). Early word learning. Handbook of child psychology, 2, 371-420.

CHAPTER 4:

Scalar implicature in absence of epistemic reasoning? The case of Autism Spectrum Disorders

Abstract

Two recent studies have argued that high-functioning individuals with autism spectrum disorders (ASD) can compute scalar implicatures. This conclusion is somewhat surprising, because scalar implicatures are thought to involve significant epistemic reasoning about speakers' knowledge and intentions, and because individuals with ASD exhibit deficits with Theory of Mind and pragmatic aspects of language. Our study explored whether individuals with ASD actually engage in epistemic reasoning when computing scalar implicatures. First, we tested whether high-functioning adolescents with ASD between the ages of 12 and 18 can compute a pragmatic inference known as "ignorance implicature." Unlike scalar implicatures, which could theoretically be computed in the absence of epistemic reasoning, ignorance implicatures necessarily require epistemic reasoning, as they involve making an inference about the mental state of a speaker. And, by some accounts, ignorance implicatures involve the same reasoning about speaker knowledge and utterance informativeness as that involved in scalar implicature. In our study, we found that subjects successfully computed both ignorance and scalar implicatures, suggesting that high-functioning adolescents with ASD have sufficient epistemic reasoning abilities to compute basic Gricean inferences in language despite their other pragmatic deficits. However, data from a follow-up task suggested that high-functioning adolescents with ASD may not always deploy this epistemic reasoning

ability in the service of scalar implicature. Specifically, we found that subjects computed implicatures both when 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 finding suggests that our ASD subjects failed to consider the specific mental states of speakers in this task. We discuss possible explanations for their failure on this task. Finally, we show that younger subjects with ASD (between 5 and 10 years of age) fail to compute both ignorance and scalar implicatures, suggesting that the ability to make Gricean inferences in language is somewhat delayed in ASD relative to typically-developing controls.

INTRODUCTION

In everyday communication, listeners 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 cause 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 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.[11]

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., $\neg B(p)$.

So, in example (1), we assume that the speaker does not believe the alternative statement in (2) to be true. 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., $B(\neg p)$).

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

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

step", that the speaker is knowledgeable regarding the truth of the stronger statement, and thus believes it to be false (Sauerland, 2004).

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. In early experimental work (Noveck, 2001; Papafragou & Musolino, 2003; Smith, 1980), it was found that children often fail to compute scalar implicatures until relatively late in development. This result was taken as evidence that implicature depends critically on pragmatic competence, since children are known to experience difficulties with other forms of epistemic reasoning early in development (e.g., false belief; Wimmer & Perner, 1983; Wellman et al., 2001). However, increasing evidence suggests that children have many of the epistemic reasoning abilities required

by scalar implicature before their failures with scalar implicature subside, sometime after the age of 6 or 7. For example, even young infants demonstrate sensitivity to speaker knowledge and intentions (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). Finally, recent studies have argued that children's difficulty with scalar implicature may not lie so much with epistemic reasoning as with 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; Chierchia et al., 2001; Foppollo et al., 2012; Stiller, Goodman, & Frank, 2011).

Recently, researchers have taken an alternative approach to understanding the nature of implicature, by investigating another group of subjects who exhibit deficits in epistemic reasoning: i.e., individuals with an Autism Spectrum Disorder (or, ASD). ASD is a neurodevelopmental disorder characterized primarily by social, communicative, and behavioral deficits (APA, 2013). Although cognitive and linguistic abilities vary widely across the autism spectrum, pragmatic deficits in language appear to be universal in ASD (e.g., difficulty with conversational turn-taking, prosody, irony, metaphor, humor, etc.; see Minshew et al., 1995; Baron-Cohen, 1988; Eales, 1993; Happé, 1993; Ozonoff & Miller, 1996; Shriberg et al., 2001; Tager-Flusberg, Paul, & Lord, 2005). These pragmatic deficits have often been linked to an impaired Theory of Mind – that is, an impaired ability to attribute mental states such as knowledge, intentions, beliefs, and

desires to oneself and other people (Baron-Cohen et al., 1985; Baron-Cohen, 1988; 1995; Frith, 2001; Tager-Flusberg, 1999). For instance, Happé (1993) found a correlation between theory of mind abilities and the ability to interpret non-literal language like metaphors and irony.

Two recent studies have investigated how individuals with ASD compute conversational inferences. First, in a study by Pijnacker, Hagoort, Buitelaar, Teunisse, and Geurts (2009), high functioning adults with ASD were asked to evaluate a set of sentences which the authors argued would require computing a scalar implicature (drawn from previous studies by Smith, 1980, and Noveck, 2001). For example, they were asked whether sentences like "Some birds are sparrows", 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 typically-developing adult controls: both groups provided "false" judgments to the same degree for target sentences. In a similar study, Chevallier, Wilson, Happé, and Noveck (2010) presented high-functioning adolescents with ASD and typically-developing controls with under-informative *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.

One conclusion that might be drawn from these studies is that, because individuals with ASD are pragmatically impaired but nevertheless can compute scalar implicatures, these inferences must therefore not strictly *require* epistemic reasoning. That is, some people – and thus possibly all people – can compute implicatures in absence of reasoning about the knowledge states of their interlocutors. However, this conclusion is premature for several important reasons. First, the studies by Pijnacker et al. and Chevallier et al. did not actually assess scalar implicature computation. In these studies, subjects were asked to make truth value judgments about sentences – i.e., to judge them as "true" and "false" or "right" or "wrong". This method, developed by Smith (1980) to test children's comprehension of quantifiers, is problematic for assessing scalar implicature computation, since under-informative statements are literally true despite being infelicitous, and thus cannot easily be judged as "true" or "false" or "right" or "wrong" (for further discussion of the limits of this method, see Katsos & Bishop, 2011; Papafragou & Tantalou, 2004). Second, the tasks in both studies did not involve reasoning about the beliefs of other individuals in context. Instead, subjects were presented sentences without context either as text on a screen or through headphones in random sequence. In each case, it is uncertain whether any reasoning about the intentional states of "cooperative, truthful, speakers" was involved. Third, although the performance of the ASD groups in these studies did not differ significantly from that of typically-developing controls, the rate at which they judged the target sentences to be false was still relatively low: in Pijnacker et al., the ASD group rejected disjunctive

statements involving scalar implicature only 53% of the time, while in Chevallier et al. the ASD group rejected these critical statements only 48% of the time. It remains uncertain, then, whether individuals with ASD truly are able to compute scalar implicatures to the same extent as typically developing controls and, more importantly, whether they actually engage in the kind of epistemic reasoning that is thought to be required for computing such inferences.

A final problem with these studies – and perhaps the greatest – lies in the assumption that ASD subjects are incapable of computing the epistemic inferences required by scalar implicature. Much like in the case of early work on implicature in typically developing children, studies on ASD have begun with the assumption that the pragmatic deficits characteristic of ASD are sufficiently strong to preclude the types of epistemic reasoning required for implicatures, and thus that they are a suitable population for assessing the role of epistemic reasoning in implicature. However, just as in the previous studies of children, this assumption has never been directly tested, leaving open the possibility that the epistemic demands of scalar implicature are within the capacities of high functioning individuals with ASD.

In this paper, we examine the role of epistemic reasoning in scalar implicature by addressing two main questions. First, we asked whether individuals with ASD are indeed capable of computing scalar implicatures under conditions that test the felicity of utterances in communicative contexts, rather than their truth value when presented on a screen or via headphones. Second, to the extent that these individuals can compute implicatures, we asked whether this ability is mediated by reasoning about epistemic states. To answer these questions, we tested individuals with ASD on two separate but

related tasks that probed pragmatic felicity rather than truth/falsity and involved sets of objects and speakers with differing states of knowledge.

The first of these tasks tested "ignorance implicature" a form of conversational inference, which, by all accounts, involves Gricean reasoning about epistemic states. Consider, for example, the utterance in (3):

(3) Billy ate pizza or pasta for lunch.

(4) Billy ate pizza.

(5) Billy ate pasta.

From this utterance, we infer that the speaker does not know which of the two foods Billy ate. More formally, upon hearing a sentence such as the one in (3), we assume that the speaker is being cooperative and has therefore provided the appropriate amount of information. From this, 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) or believed them to be false (i.e., Step III-b). However, if the speaker believed that either statement was false, then he would not 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).

By Neo-Gricean accounts, ignorance and scalar implicatures involve similar computations and epistemic reasoning abilities, and in fact, as shown in Steps I – III above, ignorance implicature is believed to be required as part of computing a scalar implicature. However, as already noted, scalar implicature could theoretically be computed in the absence of any epistemic reasoning (e.g., via compensatory strategies or via a grammatical operator), whereas ignorance implicature necessarily requires mental state reasoning, as it involves making an inference specifically about the epistemic state of the speaker. Therefore, assessing ignorance implicature computation in subjects with ASD is useful to determining whether or not their success with scalar implicature depends on epistemic reasoning. We reasoned that if individuals with ASD compute scalar implicatures and do so via epistemic reasoning, then they should also succeed ignorance implicature. On the other hand, if individuals with ASD compute scalar implicatures despite an inability to reason about epistemic states, they may compute scalar implicatures despite failing to compute ignorance implicatures.

One problem with this first set of tasks, however, is that it relies on an indirect measure of epistemic reasoning in scalar implicature: if individuals with ASD are able to compute both forms of inference, this would be consistent with a situation in which they compute scalar implicature via epistemic inference. However, such results would also be consistent with a situation in which subjects compute both inferences independently – i.e., computing ignorance implicatures via epistemic reasoning, but scalar implicature via grammatical operator (e.g., Chierchia et al., 2012; Fox, 2007). Thus, our second task sought a more direct test of the role that epistemic reasoning plays in scalar implicature. This task, inspired by a study by Bergen and Grodner (2012), asked whether subjects

consider the specific epistemic states of their interlocutor when computing scalar implicature. Specifically, we tested whether individuals with ASD compute scalar implicatures in contexts where the speaker is not fully knowledgeable and, thus, where a scalar implicature is not justified. 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 the professor who graded all of the exams in the class. This professor knows exactly how many students passed; thus, she chooses to utter the weaker statement in (6) because she knows the stronger alternative statement, "All of the students passed the exam" to be 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; he only knows that a few did. He therefore chooses an utterance with the weaker term, *some* because he 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. In our *Speaker Knowledge* task, 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.

Taken together, these two experiments allowed us to determine whether high-functioning adolescents with ASD (1) compute scalar implicatures in communicative contexts, (2) are able to compute epistemic inferences commonly thought to underlie scalar implicature on Gricean accounts, and (3) reason about the epistemic states of interlocutors when computing implicatures.

EXPERIMENT 1

Methods

Participants

We first tested 17 adults (8 males; *M* = 22.6; range = 18-41) to establish the validity of our methods. Then, we tested 17 adolescents with an autism spectrum disorder (13 males, *M* = 15.0, range = 12-18). All subjects reported English as their first language.

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.

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.

Table 4.1: Description of Age, IQ scores, and PPVT scores for the ASD group in Experiment 1

	Mean	**Range**	**Standard Deviation**
Age	15.0	12-18	1.9
IQ	96.8	71-132	19.4
PPVT[12]	100.7	56-130	22.9

Design

Subjects were tested on three different tasks, in the following order: (1) a *Disjunctive Implicature* task that included an *Ignorance Implicature* sub-task and a *Scalar Implicature* sub-task (both modeled after Hochstein et al., under review), (2) a *Speaker Knowledge* task, and (3) a *2-scales* task (not reported in this study). Once these

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

tasks were completed, subjects were also administered the Peabody Picture Vocabulary test (PPVT; Dunn & Dunn, 1997), a measure of vocabulary, and the Strange Stories Test (Happé, 1994), a measure of Theory of Mind.

The PPVT (Dunn & Dunn, 1997) is 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.

The Strange Stories Test 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, and 2) whether they were correct or incorrect.

Disjunctive Implicature Tasks

Subjects were seated at a small table directly across from the experimenter and were administered two sub-tasks: the *Ignorance Implicature* and *Scalar Implicature* tasks. In each task, subjects shown 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 sub-tasks was counterbalanced across subjects.

Ignorance Implicature Sub-Task

In the *Ignorance Implicature* 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.

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 4.2 below, and these trials were presented in one of two counter-balanced orders.

Table 4.2: 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.	Blindfolded doll
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. 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 Sub-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.

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, as shown in Table 4.3 below, and these trials were presented in one of two counter-balanced orders.

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 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 4.3: 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

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, disjunctive statements typically generate an ignorance implicature, whereas disjunctive statements containing "each" do not necessarily 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 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.

Speaker Knowledge Task

In this task, 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.'"

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 4.4, below, and 3 trials for each trial type. On each trial, Farmer Brown 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 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?" After subjects responded, 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 *2* trials, Farmer Brown said, "Two of the boxes have [strawberries]." On *all* trials (which served as controls and only occurred with *Full-knowledge* trials), 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?"

Table 4.4: Speaker Knowledge Task conditions

Speaker Knowledge	**Term used**	**Trial Type**	**Farmer Brown said…**	**Correct response to query about 3rd box**
Full-knowledge *(Farmer Brown looked inside all 3 boxes)*	All	Control	"All of the boxes have [strawberries]"	'Yes'
	Some	Full-Some	"Some of the boxes have [strawberries]"	'No'
	Two	Full-2	"Two of the boxes have [strawberries]"	'No'
Partial-knowledge *(Farmer Brown only looked inside 2 boxes)*	All	----	----	----
	Some	Partial-Some	"Some of the boxes have strawberries"	'I don't know'
	2	Partial-2	"Two of the boxes have strawberries"	'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 take the epistemic step and 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 take the epistemic step, 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.

Note that the *Full-knowledge some* trials in this task required subjects to compute scalar implicatures in a way that more closely 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 tasks of scalar implicature (including the task described above), subjects do not actually have to do this; instead, they are provided with the true state of affairs in the world and are simply asked to judge the appropriateness of an utterance in describing this (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.

2.2 Results

Disjunctive Implicature task

Prior to analysis, we compared performance on 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. Figure 4.1 displays performance by ASD subjects and typically-developed adults on the critical and control trials in both tasks.

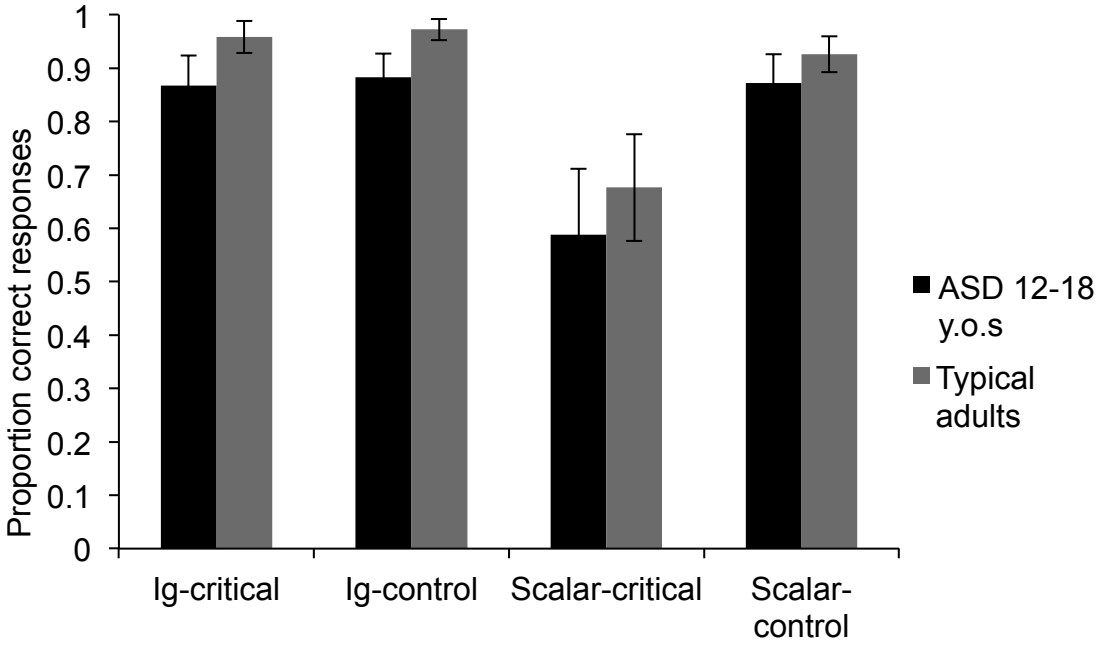

Figure 4.1: Proportion of correct responses by ASD subjects and typically-developed adults on critical and control trials in the Ignorance and Scalar Implicature tasks. Error bars represent standard error of the mean.

First, we compared our ASD subjects' performance on the *Ignorance Implicature* task to that of typically-developed 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 typically-developing adults on the *Ignorance Implicature* task.

Next, we compared our ASD subjects' performance on the *Scalar Implicature* task to that of typically-developing 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 ($z = -3.66$ $p = .00025$) suggesting that subjects performed better overall on control trials than on critical trials in this task, but 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 typically-developing adults.

To compare performance across the two tasks in just the ASD adolescent 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. This model found an interaction of Task and Trial Type ($z = -3.21$, $p = .0013$), but no other significant effects. This interaction was driven by a significant difference between critical ($M = .59$, $SE = .10$) and control ($M = .87$, $SE = .03$) trials in the *Scalar Implicature* task (Wilcoxon $T = 7$, $N = 10$, $p = .039$) but no significant difference between critical ($M = .87$, $SE = .03$)

and control (*M* = .88, *SE* = .02) trials in the *Ignorance Implicature* task (Wilcoxon T = 13, N = 7, p > .05). Thus, although the adolescents with ASD computed scalar implicatures on critical trials in the *Scalar Implicature* task to the same degree as typically-developing adults, they still did not compute them to the same extent that they responded correctly on control trials.

Finally, we explored whether there was any relation between age, IQ scores, and Strange Stories test scores and performance on critical vs. control trials in the *Ignorance Implicature* task. We used a binomial logit mixed effects model to predict binary response using Trial Type (Critical vs. Control) as a fixed variable, Age, IQ scores, and Strange Stories test scores as continuous variables, and subject as a random factor. This model found a main effect of Trial Type (z = 2.506, *p* = 0.012), a main effect of IQ (z = 2.50, *p* = 0.013), and an interaction of TrialType and IQ (z = -2.416, *p* = 0.016). This interaction was driven by the fact that there was a larger effect of IQ on critical trials in this task than on control trials. We used the same model to explore the relation between age, IQ scores, and Strange Stories test scores and performance on critical vs. control trials in the *Scalar Implicature* task. This model found a main effect of Trial Type (z = -2.244 , *p* = 0.025), indicating that subjects performed better overall on control than on critical trials, and a marginally significant interaction of Trial Type and IQ (z = 1.727, *p* = 0.084), but no other significant effects.

Speaker Knowledge task

Figure 4.2 displays performance by ASD subjects and typically-developed 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.

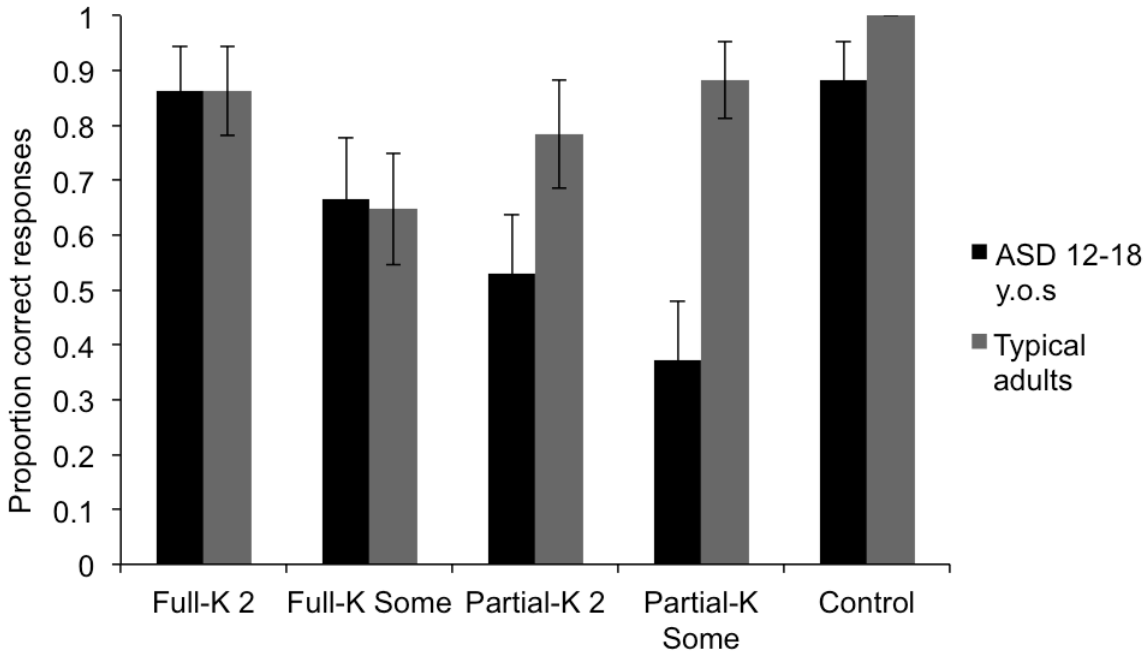

Figure 4.2: Proportion of correct responses by ASD subjects and typically-developing adults on the Speaker Knowledge task. Error bars represent standard error of the mean.

First, we explored typically-developed adults' performance on this task to establish a baseline for evaluating the ASD group's performance. We used a binomial logit mixed effects model to predict Response (correct vs. incorrect) using Condition (Partial Knowledge vs. Full Knowledge) and Term (Some vs. Two) as fixed variables and subject as a random factor. This model found a main effect of Condition ($z = 3.089$, $p = 0.0020$), indicating that subjects performed better overall on Partial-Knowledge trials

than on Full-Knowledge trials, a main effect of Term (z = 2.85, *p* = .0044), suggesting that subjects performed better overall on *Two* trials than on *Some* trials, and an interaction of Condition and Term (z = -3.0, *p* = .0027). This interaction was driven by a significant difference between *Some* trials (*M* = .65, *SE* = .10) and *Two* trials (*M* = .86, *SE* = .08) trials in the *Full-Knowledge* condition (Wilcoxon T = 5, N = 8, *p* > .05) but no significant difference between *Some* trials (*M* = .88, *SE* = .07) and *Two* trials (*M* = .78, *SE* = .09) trials in the *Partial Knowledge* condition (Wilcoxon T = 3.5, N = 6, p > .05).

Then, we explored performance on this task in the ASD group. We used a binomial logit mixed effects model to predict Response (correct vs. incorrect) using Condition (Partial Knowledge vs. Full Knowledge) and Term (Some vs. Two) as fixed variables and subject as a random factor. This model found a main effect of Condition (z = -4.0, *p* < .001), indicating that subjects performed better overall on the *Full-Knowledge* trials than on the *Partial-Knowledge* trials, and a main effect of Term (z = 2.9, *p* = .0038), indicating that subjects performed better overall on *Two* trials than on *Some* trials, but no significant interaction.

Next, we compared our ASD subjects' performance on the *Speaker Knowledge* task to that of typically-developing 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 main effect of condition (z = .3.96, p < .0001), suggesting that subjects performed better overall on the Full-Knowledge trials, a main effect of Term (z = 2.90, p = .0037), suggesting that subjects performed better overall on "two" trials than on "some" trials, and an interaction between

Group and Condition (z = 5.08, p < .0001). This interaction was driven by a significant difference between Full-Knowledge (*M* = .76, *SE* = .085) and Partial-Knowledge trials (*M* = .45, *SE* = .099) 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 typically-developing adult group (Wilcoxon T = 24, N = 11, *p* = .43).

Finally, we explored whether there was any relation between age, IQ scores, and Strange Stories test scores and performance on Full-Knowledge vs. Partial-Knowledge trials in this task in the ASD group. We used a binomial logit mixed effects model to predict binary response using Condition (Full-Knowledge vs. Partial Knowledge) as a fixed variable, Age, IQ scores, and Strange Stories test scores as continuous variables, and subject as a random factor. This model found a significant interaction of Condition and IQ (z = 1.97, *p* = .049), which was driven by the fact that there was only a marginally significant effect of IQ in the *Full-Knowledge* condition (z = 1.901, *p* = 0.06), but there was a highly significant effect of IQ in the *Partial-Knowledge* condition (z = 4.26, *p* < .000).

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 typically-developing adults, they failed to cancel these inferences on Partial-Knowledge trials.

Discussion

In Experiment 1, we sought to answer three main questions about pragmatic inferences in high-functioning individuals with ASD. The first question was whether high-functioning adolescents with ASD can compute scalar implicatures (as has been claimed by recent studies), and we found evidence suggesting that they can. In the *Scalar Implicature* task, subjects computed scalar implicatures in disjunctive statements 59% of the time, and their performance on this task did not differ significantly from that of typically-developing adults. Their performance on these trials was also slightly higher than that in previous studies (e.g., in Pijnacker et al., 2009 they computed scalar implicatures 53% of the time, and in Chevallier et al., 49% of the time). This finding was further corroborated in the *Speaker Knowledge* task, where ASD subjects computed scalar implicatures in statements containing the term *some* about 67% of the time on the *Full-knowledge* trials, and their performance on these trials did not differ from that of typically-developing adults. As mentioned in the *Results* section, the *Speaker Knowledge* task provided a more naturalistic test of scalar implicature computation than the *Scalar Implicature* task and previous studies of scalar implicature in ASD. Instead of asking subjects to judge the appropriateness of an utterance given the true state of affairs (in a picture or in a scene), the *Knowledge* task asked subjects to infer the true state of affairs on the basis of a given utterance. Thus subjects' relatively high performance on these trials in the *Speaker Knowledge* task provides especially strong evidence that high-functioning individuals with ASD can compute scalar implicatures.

Our second question was whether high-functioning adolescents with ASD are capable of engaging in the kind of epistemic reasoning that is, by some accounts, required

for scalar implicature. To this end, we tested subjects on ignorance implicature – an inference which necessarily requires mental state reasoning, as it involves making an inference about the epistemic state of a speaker. ASD subjects performed very well on this task – better even than on scalar implicature – and computed ignorance implicatures 87% of the time, which did not differ significantly from their performance on control trials. Their performance on critical vs. control trials on the *Ignorance* task also did not differ from that of typically-developed adults. These results suggest that our subjects had sufficient pragmatic understanding to go beyond the literal meaning of an utterance by reasoning about the relation between speaker knowledge and utterance informativeness. Thus, although high-functioning adolescents with ASD may have pragmatic deficits in other domains (e.g., interpreting sarcasm or prosody), they have sufficient epistemic reasoning abilities to compute basic inferences in language.

Our final aim in Experiment 1 was to test the role of epistemic reasoning in scalar implicature more directly. Although the results from the *Ignorance Implicature* task suggest that our subjects were capable of basic Gricean reasoning, it did not address whether they actually engaged in this kind of reasoning when computing scalar implicatures. It is possible, for example, that individuals with ASD compute ignorance implicatures and scalar implicatures via entirely different mechanisms. This is, in fact, the view proposed by grammatical accounts of scalar implicature, which claim that ignorance implicatures are computed via Gricean epistemic reasoning but scalar implicatures are computed via grammatical operator. We therefore designed the *Speaker Knowledge* task to examine whether subjects generate scalar implicatures when given evidence that the speaker had full knowledge about the true state of affairs but not when

given evidence that the speaker only had partial knowledge. Our subjects generated scalar implicatures on both the *Full-Knowledge* and *Partial-Knowledge* trials, suggesting that they failed to consider the speaker's epistemic state when computing scalar implicatures.

Given that subjects had little difficulty reasoning about speaker knowledge and utterance informativeness in the *Ignorance Implicature* task, it is somewhat surprising that they failed to take speaker knowledge into account in the *Speaker Knowledge* task. There are a few possible explanations for this finding. First, it is possible that although high-functioning adolescents with ASD are capable of the kind of epistemic reasoning that, by Gricean accounts, is required by scalar implicature, they do not actually compute scalar implicatures in a Gricean way – either because scalar implicatures do not generally involve epistemic reasoning (e.g., Chierchia et al., 2001) or because individuals with ASD have learned to compute these inferences via a different mechanism than typically-developing children (and via a different mechanism than ignorance implicature).

A second possibility is that high-functioning individuals with ASD do not generally reason about the specific epistemic states of speakers when interpreting utterances, but they succeeded on the epistemic reasoning component of the *Ignorance Implicature* task simply because this task forced them to engage in this kind of reasoning. The goal of the *Ignorance Implicature* task was to decide whether it was the seeing or blindfolded doll that uttered a given statement; subjects were therefore forced to make a decision about the epistemic state of the speaker in order to complete the task. In contrast, the goal of the *Speaker Knowledge* task was simply to guess the contents of a box. Thus, although our subjects may have been able to reason about specific speaker's epistemic states when explicitly asked to do so in the *Ignorance Implicature* task, they may have

failed to engage in this kind of reasoning spontaneously in the *Speaker Knowledge* task. Instead, they may generally compute implicatures by assuming that all speakers are fully-knowledgeable.

Finally, it is also possible that the *Speaker Knowledge* task was simply more difficult than the *Ignorance Implicature* task. This task required subjects to keep track of the contents of the visible boxes, to remember exactly what Farmer Brown saw or did not see on any given trial, and to integrate all this information with Farmer Brown's utterance to make a guess about the contents of a box. However, it was only on the *Partial-Knowledge* trials that the ASD group performed differently from the typical adult group; on control trials and on *Full-Knowledge* trials, there was no significant difference between the ASD and typical adult group. This suggests that it was not the task overall but instead something about the *Partial-Knowledge* trials specifically that the ASD group had difficulty with. And indeed, there was a highly significant effect of IQ on performance in the *Partial-Knowledge* trials but not in the *Full-Knowledge* trials. A correct response on the *Partial-Knowledge* trials on the *Knowledge* task required saying, "I don't know" to a yes-or-no question. Though subjects were told that "I don't know" was an acceptable response, they may have still been reluctant to use this option – either because the answers "yes" and "no" were more salient in the context of the question or because they thought that "I don't know" indicated ignorance or failure to complete the task. Thus, although the results from the *Knowledge* task raise the interesting possibility that individuals with ASD compute scalar implicatures without engaging in epistemic reasoning, more research is necessary to determine whether this is actually the case.

The results from Experiment 1 indicate that, as a group, high functioning 12- to 18-year-olds with ASD are generally able to compute both scalar and ignorance implicature. However, typically-developing children generally succeed at ignorance inferences much earlier – around 5 years of age (Hochstein et al., under review). Thus, implicature computation in individuals with ASD may still be delayed relative to typically-developing controls. In Experiment 2, we tested younger subjects with ASD on the *Ignorance Implicature* task from Experiment 1 to determine whether individuals with ASD exhibit a developmental delay in computing implicatures.

EXPERIMENT 2

Methods

Participants

Twelve children with an autism spectrum disorder (10 males, $M = 8.08$, range = 5-10) were recruited for this study. All subjects reported English as their first language. 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 criteria (APA, 2004) 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 2 or Module 3 of the ADOS, depending on their age and maturity level.

Intelligence was assessed using the Wechsler Intelligence Scale for Children (WISC) or the Wechsler Abbreviated Scale of Intelligence (WASI; Weschler, 1939).

Because the WASI can only be administered to subjects 6 years of age and older, subjects under 5 years of age in our study were not tested on a measure of IQ.

Finally, none of the subjects had any other known neurological or genetic conditions or significant hearing, visual, or physical impairments.

Table 4.5: Description of Age, IQ scores, and PPVT scores for the ASD group in Experiment 2

	Mean	**Range**	**Standard Deviation**
Age	8.1	5-10	1.83
IQ[13]	108.4	81-130	15.03
PPVT[14]	102.2	56-136	22.95

Design

Subjects were tested on the *Disjunctive Implicature* tasks described in Experiment 1. Once this was completed, subjects were also administered the Peabody Picture Vocabulary test (PPVT; Dunn & Dunn, 1997), a measure of vocabulary, and a battery of Theory of Mind assessments based on Wellman and Liu (2004).

Results

Prior to analysis, we compared performance on control trials within each task and found no significant differences between them. Critically, for all analyses reported below,

[13] Two subjects were 5 years of age and thus too young for the WASI test. These subjects were therefore not included in any analyses regarding IQ.
[14] One subject did not complete the PPVT. This subject was therefore not included in any analyses regarding PPVT scores.

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. Figure 4.3 displays performance on the critical and control trials in both tasks by both the younger ASD group in Experiment 2 and the older ASD group from Experiment 1.

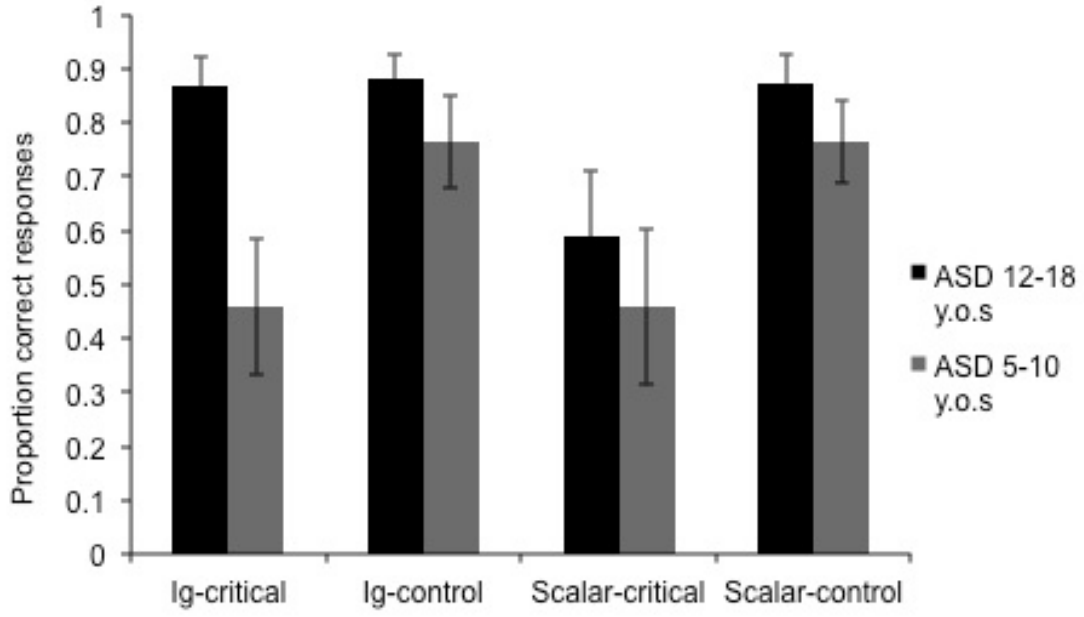

Figure 4.3: Proportion of correct responses on critical and control trials in the Ignorance and Scalar Implicature tasks by ASD age group. Error bars represent standard error of the mean.

To compare performance across the two tasks, we used a binomial logit mixed effects 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. This model found a main effect of Trial Type (z = -3.83,

p = .00013), indicating that subjects performed significantly better on control trials than on critical trials in both tasks, but no other significant effects. Thus, the 5-10-year-olds with ASD in this Experiment did not perform differently on the *Ignorance* and *Scalar* tasks, and performed significantly better on control trials than on critical trials in each task. They therefore looked somewhat similar to typically developing 4-yr-olds, who fail at critical trials on both *Scalar* and *Ignorance Implicature* tasks despite succeeding on control trials (Hochstein et al., under review).

To compare performance by the younger group in this Experiment to the older group from Experiment 1 on the *Ignorance Implicature* sub-task, we used a binomial logit mixed effects model to predict binary response (correct vs. incorrect) using Age (children vs. teens) and Trial Type (critical vs. control) as fixed variables, and subject as a random factor. This model found a main effect of Trial Type (z = -3.97, $p < .0001$), indicating that subjects performed better overall on control trials than on critical trials, and an interaction of Age and Trial Type (z = 2.87, p = .0042). This interaction was driven by a significant difference between critical (M = .46, SE = .13) and control trials (M = .76, SE = .086) in the younger child group (Wilcoxon T = 2, N = 7, $p < .05$), but no significant difference between critical (M = .87, SE = .009) and control trials (M = .96, SE = .016) in the older teenage group (Wilcoxon T = 13, N = 7, $p > .05$). Thus, whereas the older teenage group from Experiment 1 correctly attributed disjunctive statements involving ignorance implicature to the blindfolded puppet and did so to the same degree that they responded correctly to control trials, the younger group did not.

To compare performance by the younger group in this Experiment to the older group from Experiment 1 on the *Scalar Implicature* sub-task, we used a binomial logit

mixed effects model to predict binary response (correct vs. incorrect) using Age (children vs. teens) and Trial Type (critical vs. control) as fixed variables, and subject as a random factor. This model found a main effect of Trial Type ($z = -3.08$, $p = .002$), indicating that subjects performed better overall on critical trials than on control trials, and no other significant effects. Thus, although the older group performed slightly better than the younger group on the critical trials in this task, this difference was relatively modest.

Finally, we explored whether there was any relation between age, IQ scores, and Theory of Mind scores and performance on critical vs. control trials in the *Ignorance Implicature* task by the younger ASD group. We used a binomial logit mixed effects model to predict binary response using Trial Type (Critical vs. Control) as a fixed variable, Age, IQ scores, and Theory of Mind scores as continuous variables, and subject as a random factor. This model found no significant effects.

We also used a similar model to explore the relation between age, IQ scores, PPVT scores, and Strange Stories test scores and performance on critical vs. control trials in the *Scalar Implicature* task. This model found no significant effects.

Discussion

In Experiment 1, we found that high-functioning adolescents with ASD between the ages of 12 and 18 successfully compute both ignorance and scalar implicature. On the basis of this finding, we have argued that high-functioning adolescents with ASD have sufficient epistemic reasoning ability to make basic Gricean inferences in language. In Experiment 2, we explored when this ability emerges in children with ASD, and, more specifically, whether it is delayed relative to typically-developing children. To this end,

we tested 5-10-year-olds with ASD on the *Ignorance Implicature* task from Experiment 1. We found that this younger group failed to compute ignorance implicatures despite succeeding at control trials. Since typically-developing children generally succeed at ignorance implicature around 5 years of age (Hochstein et al., under review), the finding from Experiment 2 suggests that the ability to make Gricean inferences in language is somewhat delayed in ASD.

It is harder to interpret, on the basis of these findings, whether the ability to compute scalar implicatures is delayed in individuals with ASD relative to typically-developing controls. We did not find a significant interaction between Age and Trial Type in the *Scalar Implicature* task, suggesting that the difference between the older and younger groups on the critical trials in this task was relatively modest. However, the mean age of our subjects in Experiment 2 was 8.1, and it remains a matter of significant debate when typically-developing children succeed at these inferences. Some researchers have claimed that children fail to compute scalar implicatures until as late as 9 years of age (e.g., Noveck, 2001), while others argue that they succeed several years earlier (e.g., Papafragou & Tantalou, 2004). To determine whether scalar implicature is delayed in ASD, then, future studies will have to directly compare typically-developing children and children with ASD on this inference, using the same methods.

GENERAL DISCUSSION

This paper explored implicature computation in children and adolescents with ASD to determine, first, whether individuals with ASD are capable of the epistemic

reasoning required to compute basic Gricean inferences in language and, second, whether they actually engage in this kind of reasoning to compute scalar implicatures.

In Experiment 1, we found that high-functioning adolescents with ASD between the ages of 12 and 18 successfully compute both ignorance and scalar implicatures. This finding is important, because it indicates that high-functioning adolescents with ASD are capable of the kind of epistemic reasoning that is required to compute basic inferences in language, despite their other pragmatic deficits. Thus, it is a mistake to assume that just because individuals with ASD exhibit *some* pragmatic deficits, they will necessarily fail at *any* pragmatic aspects of language. And, similarly, it is highly problematic to use ASD as a general test case for determining whether pragmatic factors are at play in specific aspects of language. For instance, 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 (more specifically, in a word learning inference known as Mutual Exclusivity). 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: "The fact that many children with ASD are able to build substantial vocabularies despite impoverished social pragmatic skills provides a preliminary suggestion that pragmatic skills may not be a necessary condition for word learning and vocabulary development" (11). However, just as the adolescents with ASD in our study were capable of the kind of epistemic reasoning necessary to compute ignorance implicatures, it is also possible that children with ASD are capable of the kind of pragmatic reasoning that is involved in word learning (by some accounts), despite their other pragmatic deficits. Before using ASD as a test for the presence or absence of

pragmatic/epistemic reasoning, researchers should be careful to characterize the pragmatic deficits in ASD more precisely.

In Experiment 1, we also found that adolescents with ASD do not show sensitivity to speakers' specific epistemic states when computing scalar implicatures – generating these inferences for both fully-knowledgeable speakers and only partially-knowledgeable speakers. These results raise the possibility that individuals ASD do not compute scalar implicatures via reasoning about specific speakers' epistemic states despite having the ability to do so (as evidenced in the *Ignorance Implicature sub-task*). However, as we discussed, it is also possible that our subjects failed at the *Partial-knowledge* trials in the *Speaker Knowledge* task for reasons that had nothing to do with their epistemic reasoning abilities. Thus, further research is needed to determine whether individuals with ASD actually engage in epistemic reasoning when computing scalar implicatures.

Finally, in Experiment 2, we found that high-functioning children with ASD between the ages of 5 and 10 fail to compute both ignorance and scalar implicatures despite succeeding at control trials. This finding suggests that although high-functioning individuals with ASD are capable of the kind of epistemic reasoning that is involved in basic Gricean inferences in language, this ability is nevertheless somewhat delayed in ASD.

There are several possible reasons why children with ASD show delays with implicature computation relative to typically-developing children. One possibility is that their delays with implicature computation reflect general delays with epistemic reasoning and Theory of Mind relative to typically-developing children. There are several studies showing that children with ASD fail at Theory of Mind tests that typically-developing

children master at a relatively early age. For instance, Baron-Cohen et al. (1985) found that high-functioning children with ASD between the ages of 6 and 16 fail at a classic test of false-belief understanding known as the Sally-Anne test, while children with Down Syndrome and typically-developing controls succeed. And, indeed, 50% of the 5-10-year old children with ASD in our study failed on two or more of the 7 Theory of Mind measures they were administered, whereas typically-developing children generally succeed at all 7 of these measures by around 6 years of age (Wellman & Liu, 2004). Another possibility is that delays with implicature computation stem from overall language delays. One of the key diagnostic criteria for Autism is delayed onset of language development, and some children with ASD do not start producing single words until 2 years of age or later (e.g., Eisenmajer et al., 1998). Thus children with ASD may show delays relative to typically-developing children in all areas of language development, which would in turn delay their acquisition of implicature. Finally, it is also possible that delays with implicature computation in ASD reflect both Theory of Mind and language delays.

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 implicatures, and instead only cause some delays in their development. However, our findings also raise the possibility that individuals with ASD do not reason about the knowledge states of specific speakers when computing scalar implicatures. Further research is therefore necessary to determine the precise role of epistemic reasoning in scalar implicature computation in ASD.

Acknowledgments

Many thanks to Dave Barner, Karen Dobkins, Melissa McIntire, Shereen Cohen, Katherine Stavropoulos, Katie Kimura, Katie Wagner, Laura Taouk, Troy Campbell, Daisy Hong, 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.

Chapter 4, in full, is currently being prepared for submission for publication. The dissertation author was the primary investigator and author of this material.

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.

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.

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. (under review). Ignorance and Inference: Do problems with Gricean epistemic reasoning explain children's difficulties with scalar implicature?

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.

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.

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.

CHAPTER 5:

Conclusion

In this dissertation, we explored the role of epistemic reasoning in implicature computation by addressing two seeming paradoxes: first, the fact that children fail at scalar implicature until relatively late in development despite exhibiting basic epistemic reasoning abilities from an early age, and, second, the fact that individuals with Autism Spectrum Disorders have been claimed to succeed at scalar implicature despite exhibiting deficits with epistemic reasoning in other domains. We proposed several different ways in which these paradoxes could be resolved.

First, we considered the possibility that the epistemic reasoning abilities required for scalar implicature computation are more complex than the ones mastered by children at an early age yet less complex than the ones high-functioning individuals with ASD have difficulty with. Second, we considered the possibility that children's failures with scalar implicature and individuals with ASD's success with scalar implicature are unrelated to their epistemic reasoning abilities. For instance, we noted that children could fail at scalar implicature if they lacked the processing resources required to compute these inferences (e.g., Reinhart, 2004; Pouscoulous et al., 2007) or if they were unable to access the specific lexical alternatives involved (e.g., Barner et al, 2011). We also noted that individuals with ASD could succeed at scalar implicature by employing other linguistic or cognitive resources to compensate for their pragmatic deficits (e.g., Pijnacker et al, 2009). Finally, we considered the possibility that scalar implicature is computed grammatically, and thus does not actually require any epistemic reasoning (e.g., Chierchia et al., 2012).

To tease apart some of these different possibilities, we tested the abilities of typically-developing children and individuals with ASD to compute both scalar implicatures and pragmatic inferences known as ignorance implicatures. By all accounts, ignorance implicatures involve significant epistemic reasoning about speaker knowledge and utterance informativeness, and by Gricean accounts, these inferences involve almost exactly the same computations as scalar implicatures, and differ mainly only with respect to the alternatives involved. Thus, evidence that children can compute ignorance implicatures would indicate that they have sufficient epistemic reasoning ability to compute basic Gricean inferences in language, and thus that their difficulty with scalar implicature is not epistemic in nature. Similarly, evidence that individuals with ASD can compute ignorance implicature would indicate that they have sufficient epistemic reasoning ability to compute basic Gricean inferences in language despite their other pragmatic deficits.

We first investigated ignorance implicature computation in typically-developing children and provided evidence in Chapter 2 that 5-year-olds succeed at ignorance implicature yet fail at scalar implicature. On the basis of this finding, we argued that children's difficulty with scalar implicature is not epistemic in nature and, instead, is most likely due to an inability to access the specific lexical items in a given Horn scale. This finding therefore helps answer the first paradox that we sought to address in this dissertation – namely, why children fail to compute scalar implicatures despite the fact that they exhibit sophisticated epistemic reasoning abilities from an early age.

However, in Chapter 2 we also found that 4-year-olds fail to compute both ignorance and scalar implicature, suggesting that the ability to compute basic Gricean

inferences in language emerges around 5 years of age. This finding was somewhat at odds with claims in the literature that children can compute other forms of pragmatic inference such as *ad hoc* implicatures at an earlier age (e.g., Stiller et al, 2011). Thus, in chapter 3, we explored the role of epistemic reasoning in children's ability to compute *ad hoc* implicatures – inferences that are formally identical to scalar implicatures but which involve contextually-derived scales rather than Horn scales. We found that 4-year-olds successfully compute *ad hoc* implicatures despite failing to compute ignorance implicatures. Since ignorance implicatures necessarily require reasoning about a speaker's epistemic state, this finding raises the possibility that children compute *ad hoc* implicatures before they are able to engage in epistemic reasoning about speaker knowledge and utterance informativeness. As we pointed out in Chapter 3, however, more research is necessary to determine whether *ad hoc* implicature really does emerge before ignorance implicature. Future studies should also explore how the development of *ad hoc* implicature and ignorance implicature coincides with the development of more general Theory of Mind abilities such as understanding of false belief.

Taken together, the findings from Chapters 2 and 3 help paint a clearer picture of the development of epistemic reasoning and implicature computation in typically-developing children. Within the first few years of life, children develop a basic understanding of other people as intentional beings with knowledge and beliefs that can differ from their own (e.g., Liszkowski, et al., 2008; Onishi & Baillargeon, 2005). This understanding is necessary but not sufficient to compute basic Gricean inferences in language; children also need to learn how speakers' knowledge, intentions, and beliefs affect their specific linguistic choices (e.g., utterance informativeness), and this ability

appears to take several more years to develop, as children do not successfully compute ignorance implicatures until around 5 years of age. The inferences children make in language before they are able to reason in a truly Gricean way (e.g., *ad hoc* implicature) may therefore not actually involve epistemic reasoning. And even once children have sufficient epistemic reasoning abilities to compute basic inferences in language, they may still lack the linguistic knowledge (e.g., knowledge of Horn scales) required to make specific inferences such as scalar implicature.

In Chapter 4, we turned our attention to the second paradox regarding scalar implicature computation in ASD. We showed that high-functioning adolescents with ASD between the ages of 12 and 18 successfully compute both scalar and ignorance implicature, indicating that they are able to reason about speaker knowledge and utterance informativeness to make basic inferences in language despite their other pragmatic deficits. However, this ability nevertheless appears to be somewhat delayed in ASD, since high-functioning children with ASD between the ages of 5 and 10 fail to compute both types of inferences. Taken together, these findings help explain the second paradox – i.e., why high-functioning adolescents and adults with ASD succeed at scalar implicature despite their other pragmatic deficits. They 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 implicatures, and instead only cause some delays in their development.

These findings underscore the fact that epistemic reasoning abilities and pragmatic abilities in language are non-monolithic, and thus that failure with some

pragmatic aspects of language or on some tests of Theory of Mind does not automatically entail failure with all other pragmatic aspects of language. More specifically, the fact that high-functioning individuals with ASD exhibit deficits interpreting ambiguity, understanding humor, irony, and metaphor, or using pragmatic stress appropriately (e.g., Ozonoff & Miller, 1996; Shriberg et al., 2001; Minshew et al., 1995) does not automatically entail that they will fail at other pragmatic aspects of language, especially since we do not know precisely what causes their difficulties with humor, irony, metaphor, etc. Similarly, just because high-functioning individuals with ASD perform worse than typically-developing controls on tests of Theory of Mind that require them to infer emotions on the basis of visual cues (e.g., the Reading the Mind in the Eyes task; Baron-Cohen et al., 1997) or to reason about speakers' intentions in uttering jokes, white lies, sarcastic remarks, and other nonliteral uses of language (e.g., the Strange Stories test; Happé, 1994), does not mean that they are completely unable to reason about other people's knowledge, intentions, and beliefs. Researchers should therefore be more cautious about using ASD as a test case for determining whether pragmatic factors are at play in various aspects of language.

Finally, although the finding that high-functioning individuals with ASD compute ignorance implicatures provides evidence that they can engage in the kind of epistemic reasoning that, by Gricean accounts, is necessary to compute scalar implicatures, this finding does not address whether individuals with ASD actually engage in such reasoning when computing scalar implicatures. Thus, in Chapter 4, we examined the role of epistemic reasoning in scalar implicature more directly and found that high-functioning adolescents with ASD compute these inferences even when given evidence that the

speaker does not have knowledge about the truth of stronger alternative utterances (i.e., in contexts where, by Gricean accounts, these implicatures should be canceled). This finding has raised the possibility that although high-functioning adolescents with ASD are capable of reasoning about a specific speaker's mental states to compute implicatures (as evidenced by their performance on the ignorance implicature task), they do not always do so when computing scalar implicatures – perhaps because individuals with ASD only consider specific speakers' mental states when explicitly asked to do so, or because scalar implicatures (unlike ignorance implicatures) do not generally involve epistemic reasoning. However, it is also possible that subjects failed at the *Speaker Knowledge* task for reasons that had nothing to do with their epistemic reasoning abilities – i.e., because this task was too difficult, or because subjects were reluctant to say, "I don't know" on the critical trials. Thus, more research is needed to determine the precise role of epistemic reasoning in scalar implicature computation in individuals with ASD.

It is also worth noting that ASD refers to a spectrum of disorders, in which cognitive, linguistic, and social abilities vary widely. We investigated epistemic reasoning abilities in only a relatively small subset of individuals with ASD – those with normal cognitive abilities and relatively intact verbal abilities. More research is needed to determine whether epistemic reasoning abilities vary across the autism spectrum, and, if so, how.

In addition to focusing on only a small subset of individuals with ASD, this dissertation also focused only on a small subset of pragmatic inferences – i.e., those derived from Grice's Maxims of Quantity and Quality. Future studies should investigate whether individuals with ASD can compute other types of Gricean inferences in

language, such as inferences derived from the Maxims of Manner or Relevance. There is some indication that these types of inferences may pose more difficulty for individuals with ASD, as Surian et al., (1996) found that high-functioning children with ASD were less sensitive to violations of the Gricean maxims of Quality and Relation than typically-developing controls and children with Specific Language Impairment.

Implications for theories of implicature

The results from this dissertation also have some important implications for theories of scalar implicature. First, in Chapter 2, we found evidence suggesting that children do not consider disjunctive alternatives as relevant to computing scalar implicatures even though they consider them relevant when computing ignorance implicatures. For instance, although the 5-year-olds in our study computed ignorance implicatures for simple disjunctive utterances like, "The bunny took a cup or a plate," which involves considering the individual disjuncts "The bunny took a cup" and "The bunny took a plate" as alternatives, they did not compute either ignorance or scalar implicatures for disjunctive utterances with a universal quantifier like, "Each animal has an apple or a strawberry." This indicates that they did not consider the disjuncts, "Each animal has an apple" and "Each animal has a strawberry" as relevant alternatives in this context. Neo-Gricean theories of scalar implicature cannot currently account for this finding without some additional assumptions or modifications. We have considered the possibility that children and adults exclude disjunctive alternatives in statements under a universal quantifier because their negation can result in inferences that are too strong or

lead to contradictions – however, this proposal is speculative and requires further research.

Similarly, neither Neo-Gricean nor grammatical accounts of scalar implicature can easily account for the finding in Chapter 3 that children succeed at *ad hoc* implicatures before they succeed at both ignorance and scalar implicatures. By Gricean accounts, *ad hoc* scalar implicature involves the same reasoning about speaker knowledge and intentionality as ignorance implicature, and both inferences involve alternatives that are contextually salient. It is thus hard to see, on this account, why children would succeed at one before the other. And if, by grammatical accounts, *ad hoc* implicatures are instead computed via grammatical operator, it remains unclear how this operator would be triggered in contextually-defined scales rather than lexical scales.

Finally, although more research is needed to determine how robust these findings are, the fact that children seem to compute *ad hoc* scalar implicature before ignorance implicature and the fact that adolescents with ASD do not seem take speaker knowledge into account when computing scalar implicatures together seem to suggest that scalar implicature does not involve epistemic reasoning about specific speakers' epistemic states – at least not in typically-developing children and in individuals with ASD. This idea is consistent with grammatical accounts, which posit a separate mechanism for the computation of scalar implicatures that does not involve epistemic reasoning, or with Levinson (2000)'s proposal that scalar implicatures are computed automatically. However, we also found some evidence that typically-developing adults take speaker knowledge into account when computing scalar implicatures, as adult controls canceled implicatures on partial-knowledge trials 88% of the time in the *Speaker Knowledge* task

in Chapter 4. These results are consistent with Bergen & Grodner (2012)'s finding that adults take speaker knowledge into account during the online interpretation of scalar implicatures, and they indicate that typically-developing adults do reason in a Gricean way when computing implicatures. Taken together, the findings from these three groups raise the possibility that although scalar implicatures generally involve epistemic reasoning, this kind of reasoning is not actually *required* to generate these inferences.

Beyond implicature

In this dissertation, we considered two seeming paradoxes about epistemic reasoning and implicature computation in typically-developing children and in individuals with ASD, and we showed how these paradoxes could be resolved with a more precise characterization of the various factors at play in implicature computation and a closer investigation of the epistemic reasoning abilities of these two groups. In the case of typically-developing children, specifically, we argued that they may have an early understanding of other people as intentional beings that can share or lack knowledge yet still struggle to understand how speaker knowledge is related to utterance informativeness. This gap between early epistemic reasoning and the ability to apply this reasoning to specific distinctions in language may also help explain children's late acquisition of other aspects of language that involve mental state reasoning.

For instance, researchers have explored the relation between epistemic reasoning (or Theory of Mind more generally) and the acquisition of mental state verbs like *think, know,* and *believe*. These verbs are mastered relatively late (around 3 or 4 years of age), which has led some researchers to suggest that children do not yet understand the

concepts necessary for mental state verb comprehension (e.g. Bartsch & Wellman, 1995; Gopnik & Meltzoff, 1997). However, just as with scalar implicature, the acquisition of mental state verbs does not only require a basic understanding of mental states; it also involves significant linguistic knowledge. Mental state verbs occur with embedded complements such as that-clauses (e.g., Claire thinks that Doug stole her computer) and if-clauses (e.g., Claire wonders if Doug stole her computer), which are considered more linguistically complex than the noun phrase complements of transitive or ditransitive verbs. Furthermore, mental state verbs are often polysemous and occur in a variety of contexts (e.g., Booth & Hall, 1995; Naigles, 2000). Thus, children's late acquisition of mental state verbs may result from their linguistic complexity, or from a difficulty mapping mental state concepts to their linguistic counterparts (e.g., Papafragou et al., 2007).

Researchers have also debated the role of epistemic reasoning in children's acquisition of definite and indefinite articles in English. To use definite and indefinite articles felicitously, speakers must reason about the knowledge states of their interlocutors, since the indefinite article *a* is generally used to refer to items that are not identifiable to the listener, while the definite article *the* is generally used to refer to items that are identifiable to the listener (e.g., Clark & Marshall, 1981). Several studies have shown that children between 2 and 4 years of age often infelicitously use the definite article to introduce a new discourse entity, and some researchers have proposed that this overuse of the definite article stems from an inability to consider their interlocutor's mental state (Maratsos, 1976; van Hout et al., 2010; Schaeffer & Matthewson, 2005). However, it is also possible that children are able to reason about the mental states of

their interlocutors, but simply do not yet have a complete understanding of the meanings of these articles or of their pragmatics.

Clearly, much more research is needed to determine the precise role of epistemic reasoning in scalar implicature computation as well as in mental state verb use, article use, and many other linguistic concepts – both in typically developing children and in atypical populations. However, the results from this dissertation provide some initial insight into the development of basic Gricean reasoning abilities in typically-developing children and in individuals with ASD, and show that by disentangling the various factors at play in linguistic phenomena like implicature and by recognizing that epistemic reasoning is non-monolithic and includes a wide range of abilities, we can help resolve outstanding puzzles in the field.

REFERENCES

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., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Journal of Child Psychology and Psychiatry*, 38, 813-822.

Bartsch, K. & Wellman, H. (1995). *Children Talk about the Mind.* New York: Oxford University Press.

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

Booth, James & Hall, William. (1995). Development of the Understanding of the Polysemous Meanings of the mental-state verb know. *Cognitive Development,* 10**,** 529-49

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, H. & Marshall, C. (1981). Definite reference and mutual knowledge. In: Joshi, A., Webber, B., & Sag, I., (Eds.), *Elements of discourse understanding*. Cambridge, England: Cambridge University Press, 10-63.

Gopnik, A. & Meltzoff, A. (1997). *Words, thoughts and theories.* Cambridge, MA: MIT Press.

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.

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.

Maratsos, M. (1976). The use of definite and indefinite reference in young children: an experimental study of semantic acquisition. Cambridge, MA: Cambridge University Press.

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

Naigles, Letitia. (1995). The Use of Multiple Frames in Verb Learning Via Syntactic Bootstrapping. *Cognition,* 58**,** 221-251

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255-258.

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, Anna., Cassidy, Kimberly, & Gleitman, Lila. (2007). When we think about thinking: The acquisition of belief verbs. *Cognition,* 105, 125-165.

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.

Pouscoulous, N., Noveck, I., Politzer, G., & Bastide, A. (2007). A developmental investigation of processing costs in implicature production. *Language Acquisition*, 14, 347-375.

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

Schaeffer, J., & Matthewson, L. (2005). Grammar and pragmatics in the acquisition of article systems. *Natural Language & Linguistic Theory*, 23, 53-101.

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.

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.

Surian, L., Baron-Cohen, S., & Van der Lely, H. (1996). Are children with autism deaf to Gricean Maxims? Cognitive Neuropsychiatry, 1, 55–71.

Van Hout, A., Harrigan, K., & de Villiers, J. (2010). Asymmetries in the acquisition of definite and indefinite NPs. *Lingua*, 120, 1973-1990.