5-7-2016

Wh-Question Comprehension: A Grammatical Deficit in Children with ASD

Manya Jyotishi
University of Connecticut, manya.jyotishi@uconn.edu

Wh-Question Comprehension: A Grammatical Deficit in Children with ASD

Manya Jyotishi

B.A., Hunter College of The City University of New York, 2012

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

At the

University of Connecticut

2016

APPROVAL PAGE

Masters of Science Thesis

Wh-Question Comprehension: A Grammatical Deficit in Children with ASD

Presented by

Manya Jyotishi, B.A.

Major Advisor____
Letitia R. Naigles

Associate Advisor____
Marie Coppola

Associate Advisor____
Deborah A. Fein

University of Connecticut

2016

Abstract

Wh-questions are problematic for children with ASD. Prior research has shown delays in comprehension of subject-wh and object-wh questions in children with ASD compared to TD peers. However, earlier comprehension studies may have been limited because their stimuli included unfamiliar verbs (e.g., hit, produced by a few children with ASD) and events featuring inanimate agents and patients (e.g., an apple and a flower). The current study addressed both of these issues by investigating whether familiar verbs and animate characters elicit robust comprehension of wh-questions with children with ASD. We also investigate whether early grammatical abilities predict later wh-question comprehension. One index of early grammar is word order; children who process subject-verb-object (SVO) sentences correctly should be able to successfully understand wh-questions. In a longitudinal study of language acquisition, children with ASD (14 children with ASD averaging 33 months of age, 17 TD children averaging 19 months of age) were shown wh-question videos via the intermodal preferential looking paradigm. Both groups showed comprehension of wh-questions but children with ASD showed a delay in their comprehension. Moreover, both their word order understanding and their non-verbal IQ obtained earlier significantly and independently predicted performance of children with ASD on wh-question comprehension at their last visit. We conclude that their earlier grammatical competence (i.e., word order) guides their later wh-question comprehension.

Wh-Question Comprehension: A Grammatical Deficit in Children with ASD

According to the DSM-V (American Psychiatric Association, 2013), autism spectrum disorder (ASD) is characterized as a developmental disorder with persistent deficits in social interaction and social communication, and with patterns of restricted and repetitive behaviors. Researchers have also proposed that language development is different and delayed in this population compared to those of normal children (Rutter, 1978; Mitchell et al., 2006; Charman, Drew, Baird & Baird, 2003). Because deficits in the linguistic and communicative domains are a large part of ASD a number of studies have focused on comparing which aspects of language are intact (Kjelgaard & Tager-Flusberg, 2001) and which other domains are problematic (Eigsti, Bennetto, & Dadlani, 2007; Eigsti & Bennetto, 2009; Tager-Flusberg, 1994). The current research contributes to and expands the field of language acquisition in children with ASD by investigating one complex grammatical component of language, wh-questions. Wh-questions seem challenging for children with ASD, as prior research has shown delays in both production and comprehension (Tager-Flusberg, 1994; Goodwin, Fein & Naigles, 2012). Some researchers have argued that children with ASD have particular difficulties with wh-questions because these are complex grammatical structures (Eigsti et al., 2007) while others have proposed that their impairments are more related to pragmatics (Tager-Flusberg, 1994). We examined the former possibility by investigating whether children with ASD understand subject-wh and object wh-questions during the same developmental period as their TD peers, using a paradigm that minimizes pragmatic demands. In order to further explore the grammatical-origins argument, we examined whether earlier grammatical competence predicted later wh-question comprehension.

What are Wh-questions?

Wh-question acquisition is interesting because these questions tap into both the grammatical and pragmatic aspects of language. Grammatically, a wh-question is a question that contains a wh-word (*what*, *where*, *when*, *why*, *how*), usually occurring in the beginning of the sentence (in English). Syntactically, these wh-words stand for information that is missing in the sentence. Some wh-questions can ask for a missing argument (e.g., “What did Mary buy?”) or an adjunct (e.g. “Why did she buy that?”). Furthermore, an argument wh-question can either ask for the subject of a sentence (1) or the object of the sentence (2).

(1) Subject question: Who __ likes Mary?

(2) Object question: Who does Mary like __?

A subject-wh-question follows the canonical English SVO (subject-verb-object) word order as it asks for the agent of the action, while an object–wh-question follows OSV (object-subject-verb) word order since the wh-word refers to the patient of the action. The distance between the wh-word and its gap has led researchers to propose that subject-wh-questions are easier to understand than object-wh-questions (van der Meer, van Atteveldt, Coopmans, & Philip, 2001).

Pragmatically, wh-questions serve several communicative functions. These questions are used by 2 to 3 year old children for information-seeking purposes, specifically, information that is unknown by the speaker, such as, “Who __ chases a toy mouse?” or “What are they drinking __?” (Goodwin, Fein & Naigles, 2015; Tyack & Ingram, 1977; Bloom, Merkin & Wooten, 1982). The pragmatic nature of these questions expands in older children as they ask more directive questions, “Why don’t we read this one?” or questions for conversational purposes, “How are you?” (James & Seebach, 1982).

TD children produce “what” and “where” questions by 27-29 months of age (Bloom, Merkin, & Wooten, 1982) and master the pragmatic function between 2 and 5 years (James and Seebach, 1982). In terms of spontaneous production of subject-wh and object-wh questions, Stromswold (1995) and Tyack and Ingram (1977) both concluded that object questions and subject questions are attested in production before 30 months of age (Stromswold, 1995).

The grammar vs. pragmatics debate about wh-question challenges in ASD

Production of wh-questions seems to be delayed and sparse in children with ASD. For example, during structured and free play sessions, even verbal children with ASD requested less information compared to their TD peers and used fewer wh-questions during natural interactions (Wetherby & Prutting, 1984; Tager-Flusberg, 1994). Su, Jin, Wan, Zhang (2014) demonstrated that Mandarin-speaking seven-year old children with ASD showed delayed understanding of some grammatical and pragmatic aspects of wh-questions compared to their age-matched TD peers and also to 12 year-old children with ASD. Goodwin et al. (2012) revealed that TD children showed stable understanding of subject-wh and object-wh questions by 32 months of age, but children with ASD did not show stable understanding until 54 months of age. These findings showed that compared to their TD peers, children with ASD showed a delayed development of wh-question comprehension. Both these studies revealed some level of grammatical deficit in comprehension of the wh-question construction using tasks that reduced pragmatic demands. However, their task demands and stimuli, respectively, might have impacted their results (discussed in detail below). These two components of wh-questions have led to a debate in the language acquisition literature concerning the extent to which the paucity of wh-question acquisition is grammatically vs. pragmatically based in children with ASD. Eigsti et al. (2007), using a fine-grained analysis of play sessions, reported that five – year old children with

ASD, matched on non-verbal IQ and receptive vocabulary with TD children and children with developmental delays, used fewer questions compared to the TD and DD groups during the play session. An examination of their pattern of responses revealed that children with ASD produced wh-questions inconsistently compared to the TD and DD group. Their question production did not follow the typical pathway, i.e., from simpler forms to more complex ones. Children with ASD followed a different developmental progression with their wh-question forms, such as using more advanced forms, "What does it do?" rather than using a simpler structure, "What's that?" Even a small measure of their syntactic ability, MLUs, revealed delays in the ASD group. This reflects their syntactic limitations, which may narrow their understanding of different grammatical constructions (for example, object-wh-questions might present a grammatical challenge to children with ASD because it follows the non-canonical OSV word order; Scarborough, Rescorla, Tager-Flusberg, Fowler, Sudhalter, 1991). These findings reflect clear syntactic deficits, specifically with wh-questions, in the ASD group.

On the other hand, Tager – Flusberg (1994) argued that the deficit in wh-question production witnessed in children with ASD is more related to pragmatics rather than grammar. The impairment lies in their social communication (Rutter, 1978), i.e., their desire to communicate with others, leading to pragmatic difficulties. Her analysis of six children with high-functioning ASD revealed that over time children with ASD showed increases in their grammatical development, with MLUs appearing in the normal range. Children with higher MLUs can possibly comprehend syntactically more complex sentences like wh-questions. In this study, children with ASD showed a gradual increase in their percentage of well-formed wh-questions. These well-formed questions included auxiliary verbs, a copula verb, or an inversion and increased for each MLU stage but the usage of these questions in conversations was more restricted and did not serve a proper conversational function of agreement and clarification that would maintain verbal interactions. Children with ASD failed to ask social routine questions like, "How are you?" or failed to use conversational openers. They asked questions that are tangential to the conversation, and sometimes they talk *to* someone rather than *with* someone (Davis, 1932). It is possible that there exists an asynchrony between grammatical components and their communicative function because MLUs of children with ASD were in the normal range as their TD peers but they showed impairments in their functional aspects of language. For example, compared to children with Down syndrome, children with ASD asked fewer questions (9.3 per 1000 utterances in ASD; 28.2 per 1000 utterances in Down syndrome). Therefore, even though children with ASD produced wh-questions, their social usage of them was deviant, indicating that the acquisition of these forms is dissociated from their function. Tager – Flusberg (1994) concluded that children with ASD's social deficits influence different aspects of their social communication.

Overall, comparing the 'grammatical' vs 'pragmatic' deficit arguments, is it the case that children with ASD use less varied and inconsistent wh-question structures because they are socially impaired in acknowledging the varied wh-question functions or because the formation of wh-questions and an understanding of wh-movement are grammatically impaired? Two experimental studies addressed some of these important issues regarding wh-question comprehension in children with ASD (Su et al., 2014; Goodwin et al., 2012). As mentioned before, both comprehension studies found a delayed comprehension of wh-questions in children with ASD, however, these are not the final word as these comprehension studies may have been flawed in numerous ways, and therefore, the low success rate in comprehension may have been related to their task demands, as discussed below.

Wh-question Comprehension in TD Children and Children with ASD

In young TD children, comprehension of wh-questions, specifically 'what' and 'where' questions, occurs by 20 months of age (Seidl, Hollich, & Jusczyk, 2003). For example, Seidl et al. (2003) used a simple IPL task to investigate comprehension of object-questions, subject-questions and where-questions in toddlers. The IPL methodology makes minimal social, cognitive, and motor demands on the child participants (i.e. children just need to look at the visual stimuli) and it is a feasible way to assess children's knowledge about these questions. Their eye movement patterns revealed that by 20 months of age, toddlers are able to show reliable understanding of all three types of wh-questions, compared to 13- and 15-month olds. Their low task demands revealed that TD children relied on syntactic information to look at the answer for these questions. These findings suggest that comprehension of such questions is quite early in TD children. On the other hand, two experimental comprehension studies have been conducted with children with ASD and both studies have suggested that comprehension of wh-questions is delayed in this population (Su et al., 2014; Goodwin et al., 2012). However, it is important to report on their methodological issues in order to show that their results are not entirely conclusive.

In a comprehension study, Su et al. (2014) investigated the interpretation of wh –words in Mandarin-speaking younger and older high-functioning children with ASD. Using a question – statement task, children were asked to respond with either a *yes* for a true statement or a *no* for a false statement regarding the facts of a story, and to respond with an answer in the question condition. Young children with ASD performed significantly more poorly than the younger TD and older ASD groups. Their poor performance indicated delayed development of some aspects (grammatical, pragmatic) of wh-questions in these Mandarin learners with ASD. But, the high level of performance of the older children with ASD demonstrates eventual mastery of these forms. However, an important caveat to note is that even though this study used a paradigm that claimed to have minimized pragmatic constraints, children with ASD were still required to *answer* questions when the wh-question was asked on the laptop, which still adds a social aspect to the task. Younger children with ASD might have performed differently if they were not asked to answer questions, but rather just *look* at the correct answer.

In a different and much simpler task, Goodwin et al. (2012) examined wh-question comprehension in English-speaking TD children and children with ASD using the IPL paradigm. The IPL (intermodal preferential looking) paradigm can help in the assessment of language comprehension in young children (Golinkoff, Hirsh-Pasek, Cauley & Gordon, 1987; Golinkoff, Ma, Song & Hirsh-Pasek, 2013). This paradigm places few motor and speech demands on the participants as children just need to watch two videos on a monitor while the audio speaker in the center plays a linguistic stimulus that only matches one of the images on the screen. The IPL measures children's changes in eye-gaze, which are presumed to be guided by the accompanying language. The logic behind the paradigm is that children will look longer or designate more attention at the image that matches the linguistic sentence. This requires little to no social interaction on behalf of the participant, thus reducing the pragmatic demands. IPL has been used to investigate comprehension of a number of linguistic constructions in children with ASD, including word order (Swensen, Kelley, Fein & Naigles, 2007), noun bias (Tek, Jaffrey, Fein & Naigles, 2008), shape bias (Tek et al., 2008), specific words (Edelson, Fine & Tager – Flusberg, 2008; Venker, Eernisse, Bean, Saffran, & Ellis Weismer, 2011) and syntactic bootstrapping (Naigles, Kelty, Jaffery & Fein, 2011). Most importantly, the IPL paradigm can address the grammar vs. pragmatics debate because it reduces the social constraints on the use of wh-questions. Children are only required to *look* at the matching visual stimulus when the auditory linguistic stimulus is presented. Children are not asked to answer any question, which minimizes the social pressure on them as there is little to no social interaction during the task, and the videos last less than 5 minutes. Thus, if wh-questions are a grammatical challenge for children with ASD, then one should see poor performance in a comprehension task using the IPL because their knowledge of e.g., syntactic movement may be impaired. In contrast, if wh-questions are a pragmatic challenge for children with ASD, then one should see intact comprehension, and e.g., better comprehension than production.

Goodwin et al. (2012) used the IPL paradigm to study 18 TD children and 15 children with ASD who were matched on their language level at their first visit. Their video consisted of familiar items, such as an apple, flower, keys, and a book, engaged in hitting events (i.e., an apple hitting a flower, keys hitting a book). Following these familiarization trials, children were presented with three test trials that asked object – and subject – *what* questions and where questions while each item was displayed simultaneously, side by side. The where audio asked, “Where are the keys/book?”; the *what* – object audio was “What did the keys hit?” and the *what* – subject audio, “What hit the book?” These questions were then repeated for the keys/book events. The TD children demonstrated reliable understanding of wh-questions by 32 months of age. However, children with ASD did not show reliable comprehension until 54 months of age. So, comprehension of wh-questions was observed in both groups, but chronologically emerged much later in development in children with ASD. Developmentally, the language level at which children with ASD showed comprehension of wh-questions was similar to the language level at which the TD children first showed comprehension of wh-questions.

This study actually supports both the ‘grammar deficit’ and ‘pragmatic deficit’ arguments: The pragmatic deficit argument is supported because the use of the IPL paradigm, minimizing children’s pragmatic demands, did reveal reliable comprehension of wh-questions by the children with ASD, including comprehension emerging developmentally prior to production. The idea of comprehension preceding production highlights the fact that children’s acquisition of linguistic forms does not depend on their ability to produce them (Maratsos, 1988). However, the grammar deficit argument is also supported, because wh-question comprehension emerged chronologically later in the children with ASD, than in the TD children. However, this study had a number of limitations, as well. First, both events involved the action of hitting, yet this verb is not a common one for children with ASD. Their last visit, when the children with ASD were 54 months of age, only 53% of those children produced the verb ‘hit’, according to their CDI (Fenson et al., 1994), a measure of their language-production abilities, via parental report. In contrast, all TD children in the study had produced this verb at 32 months of age. Second, the hitting events shown involved the action of an inanimate agent on an inanimate patient; these are non-prototypical actions (Slobin, 1982) that might have caused some confusion in this atypical population. Slobin (1982) defines a prototypical transitive event as an event in which an animate agent is acting upon an inanimate patient, which was not the case in these videos as both objects were inanimate, thereby possibly leading to confusion. Therefore, it is possible that earlier comprehension of wh-questions in children with ASD was not demonstrated due to the above reasons and that changing the stimuli might tap into earlier comprehension.

The Subject-Wh vs. Object-Wh Questions Asymmetry Debate

Another aspect of wh-questions is the asymmetry in comprehension of subject-wh and object-wh questions. Subject-wh questions have been found to be easier to process than object-wh questions, probably due to the shorter wh- movement between the wh–word and its gap in subject-wh questions, and also due to the unchanged surface word order of the sentence (van der Meer et al., 2011). For example, in the sentence "Who __ is helping the boy?" the 'who' refers to the agent of the sentence whereas 'who' refers to the patient of the sentence in, "Who is the boy helping __?" It may be more difficult to process object–wh–questions since the wh–word is further removed from its gap (3). It has to cross over another argument to stand in the beginning of the sentence.

(3) Object question: Who is the boy helping__?

Furthermore, in some atypical populations, such as children and adolescents with Specific Language Impairment (SLI), object-questions have also been reported to be more impaired than subject-questions (Ebbels & van der Lely, 2001; Stavrakaki, 2006). Most of these studies, have investigated *which* and *who* wh-questions using picture-selection tasks. For example, one image shows a 'cat biting a dog' and the other image shows a 'dog biting a cat'. When children were asked *who*-subject, *who*-object and *which*-subject and *which*-object questions, 'Who is biting the cat?', 'Which dog is biting the cat?', 'Who is the cat biting?' and 'Which dog is the cat biting?' children with syntactic SLI had the most difficulties with *which*-object questions and performed significantly worse on these than the control group (Friedmann & Novogrodsky, 2011; see also studies with Broca's aphasics: Avrutin, 2000; Hickok & Avrutin, 1996; hearing impairment: Friedmann, Szterman & Haddad-Hanna, 2010; Friedmann & Szterman, 2011). However, the object-wh-question difficulty is not universal; for example, children with SLI performed at comparable levels on *who*-object and *who*-subject questions. Deevy and Leonard (2004) also observed no differences between *who*-subject vs. *who*-object questions in five-year old children with SLI (also with Cantonese-speaking children, Wong, Leonard, Fletcher, & Stokes, 2004).

With TD samples, the findings are also mixed: four-year old TD children showed significantly poor performance on long – object questions, such as, "Who is the happy brown dog washing x?" (Roeper, 2004) whereas Stromswold (1995) revealed, after analyzing transcripts of 12 children between the ages of one to six years, that children acquired object questions developmentally *before* subject questions.

In sum, the degree to which object-wh questions are indeed harder to process or more challenging to acquire is still an open question. Current studies are limited by stimuli (only Goodwin et al., 2012, have used what-questions) and method (picture-pointing tasks may elicit points to both images, and/or confusingly points to multiple figures in the images). It seems that if wh-questions pose a grammatical challenge to children with ASD, then they should demonstrate impairments in *both* subject-wh and object-wh-question comprehension as both these questions involves syntactic movement of a noun phrase. In subject-questions, the movement occurs from the subject position (4), and in object-questions, the movement occurs from the object position (5). Therefore, if children have considerable difficulty with grammatical forms, then children should demonstrate impairments in both these structures that are derived by wh-movement.

(4) Subject question: Who __ likes Mary?

(5) Object question: Who does Mary like __?

Does Early Grammatical Competence Guide Wh-Question Comprehension?

The knowledge that wh-questions involve wh-movement involves a degree of syntactic ability. An important area to investigate would be to examine how children's prior grammatical knowledge contributes to their later syntactic abilities, i.e. wh-question comprehension. A number of studies have investigated how children's processing speed of words predicts their later language and cognitive outcomes. More specifically, Fernald, Perfors, and Marchman (2006) revealed using a looking-while-listening (LWL) paradigm that online speech processing (reaction time to a word image pair) at 25 months is associated with vocabulary growth across 12, 15, 18, 21 and 25 months in English-speaking children. Moreover, children who processed speech faster at 25 months had accelerated vocabulary growth during their 3rd year. Marchman and Fernald (2008) demonstrated that processing speed of word recognition and infants' vocabulary size at 25 months predicted their cognitive skills at 8 years, using standardized tests of language, cognition and working memory. Venker, Eernisse, Saffran and Weismer (2013) examined real-time lexical processing in children with ASD using the LWL paradigm and found that children with better accuracy on familiar words processed those words faster. Their processing speed of words was also associated with their earlier vocabulary comprehension. Thus far, only one study has examined predictive relations between children's processing speech of sentences and their later sentence comprehension (Naigles et al., 2011). Seventeen children with ASD and 18 TD children were taught novel verbs in transitive sentences via the IPL paradigm and then asked whether the novel verbs matched to causative or non-causative actions. Both TD children and children with ASD were able to successfully interpret novel verbs in transitive sentences as causative actions, thereby engaging in syntactic bootstrapping (Naigles et al., 2011). This indicated that abstract syntactic knowledge is present in children with ASD. A strong predictor of syntactic bootstrapping was also their earlier performance on a word-order task. Children who were faster processors of SVO word order eight months earlier were able to use abstracted SVO frames to make predictions about new verb meaning. Their earlier SVO grammatical knowledge helped them to generalize grammatical patterns using novel verbs in transitive sentences.

Therefore, word order is an important syntactic domain to investigate, with respect to wh-question comprehension, as well, because if children have understood the subject-verb-object (SVO) frame, then they should be more likely to understand subject-wh-questions because these questions follow the same SVO pattern. For example, in order to engage in wh-movement, children should have abstracted the SVO sentence frame (6) and matched the structure of the frame with the wh-question (7) to help them guide to the correct referent (either the agent or patient) of the action.

(6) John likes Mary.
S V O

(7) Who ___ likes Mary?
S V O

(8) Mary likes John.
S V O

(9) Who does Mary like ___?
S V O

In the above example, if children have abstracted the 'agent-verb-patient' frame from hearing the sentence 'John likes Mary', then when they hear a subject-wh-question like, "Who __ likes Mary?" children would know that the *who* refers to the gap in the subject position of the question, 'Who ___ likes Mary?' This would indicate that children are able to generalize the sentence structure rule to wh-questions, indicating that a precursor to wh-question understanding might lie in their early abstraction of the SVO sentence frame, i.e., children used their early grammatical knowledge to learn about the agent or patient referent in the subject-wh or object-wh-question (later grammatical comprehension). Moreover, abstraction of the SVO sentence frame can also help with object-questions (9) if children understand that the SVO frame is a transitive frame with an agent (a '*liker*') and a verb ('*like*') that requires a direct object (a '*likee*') and that these map onto a causative action (Naigles et al., 2011). This may lead children to map the wh-word movement back to its gap in the patient position as the question is asking for the *object* of the sentence (9), given the subject and the verb. So, grammatical knowledge of wh-questions might emerge from children's level of knowledge of such word order frames. Therefore, analysis of such predictive associations can lend support to the grammar deficit argument in wh-questions. If children who have difficulties with the abstraction of the SVO frame also demonstrate difficulties with subject-wh or object-wh-question comprehension, then we can make a stronger argument for the 'grammatical' deficit viewpoint. So, in the current study, we examined the relationship between children's performance on an earlier word order IPL task and their performance months later on the wh-question comprehension measure. Thus, impairment in early grammatical abilities should contribute to later wh-question performance in children.

Current Study

In the current study, we both addressed and expanded upon the questions of Goodwin et al. (2012), i.e., do children with ASD comprehend wh-questions at the same visits as their TD peers? Because of the limitations of Goodwin et al.'s (2012) study (i.e., using videos with inanimate characters engaged in non-prototypical actions, using a verb *'hit'* that they likely did not understand), we modified the videos in the current study. It is possible that we did not see an earlier comprehension in children with ASD due to these caveats in their study. Therefore, with a new sample of children, our study used the IPL paradigm with new videos that included animate characters, i.e., a costumed horse and a bird. These videos also used new verbs, such as *tickle*, *wash*, *hug*, and *ride*, which have been reported to be understood by children with ASD at 2.5 years of age on their CDI. An examination of their CDI showed that children used these verbs at least once with ASD at visit 4. The new stimuli enable us to address the grammar vs. pragmatics debate about wh-questions because these modified videos could possibly tap into earlier wh-question comprehension; however, if we still see delayed comprehension with the new videos in the new sample of children using the IPL paradigm which is less pragmatically stressful, we can argue that the wh-question impairment has a grammatical root. Also, if a grammatical impairment does exist then children should show delays in *both* subject and object-wh-questions. Moreover, we also examined the relationships between early standardized test measures and word order comprehension and later wh-question comprehension. If early grammatical competence is associated with later performance on wh-questions, this strengthens our argument of a grammatical deficit in wh-questions in children with ASD.

Method

Participants

Fourteen children with ASD and seventeen TD children participated in this longitudinal study. All were monolingual English learners. One child with ASD participated in the overall project, but was not included in the final analyses of this study because he failed to provide sufficient data during the wh-question task for more than half of the visits. One child in the TD group was omitted from the IPL analyses at visit 6 because she had missing data at this visit. We recruited participants in the ASD group by contacting facilities that offer Applied Behavioral Analysis (ABA; Lovaas, 1987); we restricted the sample to children receiving ABA to ensure some consistency in the interventions being received. Moreover, ABA is the most common intervention offered in our geographic area (northeastern U.S.). These service providers distributed information about the study to parents of children who had been diagnosed within the last 6 months and had just begun ABA training. Interested parents then contacted us and were interviewed via telephone to verify their child's diagnosis and eligibility for the study. All parents signed consent forms prior to participating.

The participants in the ASD group included seven White males, two Asian males, and one African American male. There were two White females, one Asian female and one African American female. This sample of children somewhat reflects the prevalence of ASD in the general population; we made significant efforts to recruit non-Caucasian families. All children were from lower-to upper-middle-class families living in the Northeastern United States. At the first visit, the children with ASD ranged in age from 18 months to 42 months (*M*=32.93, *SD*=7.28). To be included in the study, the children with ASD had to be receiving at least 20 hours of ABA intervention weekly. Because it is difficult to distinguish between ASD and pervasive developmental disorder – not otherwise specified (PDD-NOS), we accepted participants with either diagnosis, which was then verified by the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 1999). The ADOS and other test scores are provided in Table 1.

The TD group was recruited via birth announcements from local newspapers. The TD group included 13 White males, three White females and one Asian female from middle- to upper-middle-class families living in Connecticut. These demographics closely resembled those of the ASD group. Rather than matching the TD group to the ASD group on age, we chose to match them on level of language development. Therefore, we began testing TD children at approximately 20 months of age (*M*=19.74, *SD*=1.25), when their language abilities were most similar to those of the ASD group at visit 1 (see Table 1).

Materials

Standardized tests. The ADOS (Lord et al., 1999) was administered to assess ASD status. We also administered the Vineland Adaptive Behavior Scales, 2nd Edition (Vineland II; Sparrow, Cicchetti, & Balla, 2005) to evaluate children's communication, socialization, daily living skills, and motor skills, which yielded standard scores. The Mullen Scales of Early Learning (1995) were administered to measure the development in the areas of visual perception, fine motor skills, receptive language, expressive language, and gross motor skills. Finally, the MacArthur Communicative Developmental Inventory (CDI; Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994) provided a measure of the child's production vocabulary, via parental report. The infant version of the CDI was used at visit 1. The Receptive One-Word Picture Vocabulary Test, 4th edition (ROWPVT-4; Martin & Brownell, 2000) and Expressive One-Word Picture Vocabulary Tests, 4th edition (EOWPVT-4; Martin & Brownell, 2000) were administered at all visits to evaluate the children's receptive and expressive vocabulary skills, respectively.

IPL setup. The IPL paradigm [Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Naigles & Tovar, 2012] involves showing children two videos side by side, while playing child–directed speech from a central speaker that corresponds to only one of the videos. The child's direction and duration of gaze are recorded and used as indications of his/her understanding. An Apple Powerbook was used to project the stimuli onto a portable 63" X 84" screen, via an LCD projector. The computer was connected to an external speaker, which was placed out of sight behind the screen. A digital camcorder for filming the child's face was placed on a small tripod in front of the screen, just below the center.

Video Stimuli

Word Order (Candan et al., 2012): The layout for the word order video is presented in Table 2. The pretest trials (labeled "P" in the table) introduced and labeled the costumed horse and bird. Trials 1-2 presented a familiar action with agent A and patient B on one side (e.g., the bird pushing the horse), and then with agent B and patient A on the other side (e.g., the horse pushing the bird). During these trials, the action was labeled in a neutral frame (e.g., "Pushing!"). In Trial 3 (the control-for-salience trial), both renditions of the action were presented simultaneously and the audio was the same as in trials 1 and 2; this provided a baseline measure of stimulus salience. Trial 4 was the test trial, in which the verb was placed in a sentence such that only one of the two renditions matched. This trial thus examined whether the child understood the difference between "A verbs B" (e.g., "the bird is pushing the horse") and "B verbs A" (e.g., "the horse is pushing the bird"). A total of six familiar verbs and actions were introduced and then tested for word order understanding. These were *push, tickle, pull, wash, hug,* and *ride*. The same characters were used for each action; the horse was the agent for half of the matching actions and the bird was the agent for the others.

Wh-Question: The wh-question video layout included familiarization trials, which introduced the video stimuli sequentially on each side of the screen; control trials, during which the two target stimuli were played simultaneously without any directing audio, to obtain baseline looking times; and test trials, during which the two stimuli were displayed side by side and the audio directed the child to look at one of them.

A costumed horse and bird served as agents and patients. They engaged in four familiar transitive actions: wash, tickle, ride and hug. At the beginning of the wh-question video, there was a baseline trial (trial 2 in Table 3). During the baseline trial, each costumed animal appeared side-by-side on the screen, and the audio prompted the child to look, without designating which side to look at. The baseline trials were followed by wh-object and wh-subject blocks, which consisted of two familiarization trials (trials 4 and 6) and a test trial (trial 8). In the familiarization trials, the transitive event was seen sequentially on each side of the screen. Then, in the test trial, a what-question was heard while each item was displayed simultaneously, side by side. The first four blocks (see Table 3; trials 1-28) asked wh-object questions (i.e. “What did the horse tickle?”). The second four blocks (trials 29-53) asked wh-subject questions (i.e. “What hugged the horse?”). The final block (trials 55-58) asked where questions (i.e. “Where is the horse/bird?”). In contrast to the other blocks, there were no familiarization trials in this block. The audio in the where-trials (trials 56 and 58) directed the child to look at the named animal. The horse and bird equally appeared as agent and patient for each event and each child was asked four object-wh-questions, four subject-wh-questions, and two where-questions. For both videos, the side of the matching scene was counterbalanced both within (i.e., the matching side varied from left to right in an XYYXXY pattern) and between (i.e., for half of the children the first match was on the left and for the other half, the first match was on the right) participants.

Procedure

The children were visited in their homes at 4-month intervals for a total of six visits. Ages at each visit are displayed in Table 4. The visits began with one experimenter administering standardized tests, while another experimenter prepared the IPL setup. The child watched three IPL videos at each visit. The word order video was shown at visits 1 and 2; the wh-question video was shown at visits 3 through 6, and was always the second or third video in the series. Breaks were allowed as needed between videos. Finally, the mother completed any remaining surveys or forms.

Coding

The films of the child's gaze during the IPL task were captured and digitized in the lab. Looking times were coded offline by watching these films frame by frame, using a custom coding program. The test audio was removed, so the coders did not know which direction of looking was correct. Looking during each frame was coded as to the left, right, center, or away. If a child did not look at either screen for more than 1 second total for a given trial, his/her data were not included for that trial. For the wh-question video, this occurred in 1.4% of test and control trials in the TD group and 4.6% of test and control trials in the ASD group. For the word order video, the percent of excluded trials in the TD group was 2.7%, and it was 2.9% in the ASD group. All participants were coded by multiple coders to ensure reliability. The correlation between coders averaged 0.99, $p < .001$.

Word-order comprehension. One dependent variable was calculated from these data; namely, the children's proportion of looking to the match during both the test and baseline trials. This is the most typical measure from IPL with dynamic scenes (Piotroski & Naigles, 2012); the test-baseline comparison demonstrates the degree to which the test audio guided the children's looking at the matching scene, relative to their initial preference for that scene based solely on stimulus salience.

Wh-question comprehension. The dependent variable was the proportion of time that the child looked at the named item during each trial type (i.e. subject-, object-, and where-questions). This was the metric employed by Seidl et al. (2003; see also Goodwin et al., 2012) to demonstrate what-question comprehension; namely, the child needed to look at the named item significantly less during a subject- or object-wh-question trial compared with during the where-question trial. For example, to assess comprehension of "What tickled the bird?" we compared children's looking time to the bird during this trial vs. during the “Where is the bird?” trial. During the “where” trial, they should look consistently at the bird whereas during the “what” trial, they should look consistently away from the bird. Similarly, comprehension of “What did the horse tickle?” was assessed by comparing children’s looking at the horse during this trial vs. the “Where is the horse?” trial. Such within-subject comparisons are common with the IPL paradigm, as children’s eye movements during baseline trials serve as their own controls for performance during test trials (Brandone et al., 2007; Piotroski & Naigles, 2012; Swingley, 2011). To succeed at this task, then, children need not demonstrate a completely adult-like understanding of the grammar, they need only to allow the ‘what’ questions to pull their attention away from the named item, indicating that they are aware that grammatical movement has occurred (and that SVO is no longer the correct word order). There is evidence that adults, too, initially look at the named item before switching to the correct referent, during online processing of what-questions (Kukona & Tabor, 2011; Sussman & Sedivy, 2003).

Results

The results of this study are organized according to two questions: (1) Did the children with ASD comprehend wh-questions at the same visits as the TD children did? (2) Did the children’s comprehension of SVO word order at visits 1 and 2 predict their later comprehension of wh-questions?

When Do Children with ASD Comprehend Wh-Questions Compared to TD Children?

The analyses used children’s percent looking time to the named item as the dependent variable, averaged separately across object-wh-test trials, subject-wh-test trials, and the two where-questions.

Preliminary pairwise comparisons were conducted to check whether children showed a preference for either the horse or the bird in the first baseline trial. Such an analysis is important to make sure that attention is equally distributed to each image and that the stimuli is neutral to children, so that when a child does show a preference during the test trial, it's because of their response to the linguistic stimulus rather than their initial preference for that figure. Across the four visits, the costumed bird and horse appeared equally salient to the TD children (visit 3: $t(15) = -0.158$, $p = .876$; visit 4: $t(16) = 0.972$, $p = .345$; visit 5: $t(16) = 1.11$, $p = .283$; visit 6: $t(15) = 0.604$, $p = .555$); see means in Figure 1a). Children with ASD looked equivalently at both the horse and the bird, except at visit 4 when they significantly preferred the bird (visit 3: $t(13) = -0.097$, $p = .924$; visit 4: $t(13) = 2.20$, $p = .046$; visit 5: $t(13) = 0.372$, $p = .716$; visit 6: $t(13) = 1.38$, $p = .190$; see means in Figure 1b). We concluded that there is no overall animal preference.

Addressing our first major question, a repeated-measures analysis of variance was conducted with group (ASD or TD) as the between-subjects variable, visit (3, 4, 5, or 6) and trial type (where-wh, object-wh, and subject-wh) as within-subjects variables. The results showed a main effect of trial ($F(2, 48) = 35.67$, $p < .001$, partial eta squared = .598), indicating that children's proportion of looking to the named object was different for object-wh-questions, subject-wh-questions, and where-questions. There was no main effect of visit ($F(3, 72) = 1.79$, $p = .157$, partial eta squared = .069), nor a significant group X trial interaction ($F(2, 48) = 389.18$, $p = .279$, partial eta squared = .052). A significant group effect emerged ($F(1, 24) = 2190.78$, $p = .02$, partial eta squared = 0.204), with greater overall looking to the named object by the TD group than by the ASD group. Given this significant group effect, the next set of analyses investigated the groups' looking patterns separately.

Figure 2 shows the TD group's percentage of looking to the named item for the where-trials compared with object-what-trials and subject-what-trials at each visit. For the purpose of these analyses, one-tailed significance testing was used as we expected an effect in a specific direction, i.e., *less* looking to the named item during the what-test trials. Children looked significantly less to the named item during the object-what-trials vs. where-trials at all visits (visit 3: $t(16) = 1.90$, $p = .038$; visit 4: $t(16) = 3.68$, $p = .001$; visit 5: $t(16) = 4.09$, $p < .001$; visit 6: $t(15) = 6.26$, $p < .001$; see Figure 2a). For subject-what questions, TD children looked significantly less at the named item during what-questions compared to where-questions starting at visit 4 (visit 3: $t(16) = 1.27$, $p = .111$; visit 4: $t(16) = 3.75$, $p < .001$; visit 5: $t(16) = 3.57$, $p = .001$; visit 6: $t(15) = 8.52$, $p < .001$; see Figure 2b).

The ASD group's performance was less consistent for object-what questions: while they appeared to show comprehension at visit 3, $t(13) = 3.39$, $p = .002$, this effect disappeared at visit 4, $t(13) = 0.998$, $p = .168$ and visit 5, $t(11) = 1.05$, $p = 0.157$, then re-emerged at visit 6, $t(11) = 2.07$, $p = .031$; see Figure 2b. Similarly, the ASD group's performance with subject-what questions varied across visits, reaching significance at visit 3 and trending towards significance at visit 5 (visit 3: $t(13) = 2.30$, $p = .019$; visit 4: $t(13) = 0.807$, $p = .217$; visit 5: $t(11) = 1.58$, $p = .07$; visit 6: $t(10) = .857$, $p = .206$; see Figure 2b).

Because each group performed similarly on the object-wh and subject wh-questions, we combined both types of what-questions and compared this to the average where-questions score to get an overall measure of wh-question comprehension. The TD group demonstrated stable comprehension of wh-questions starting at visit 4 (visit 3: $t(16) = 1.70$, $p = .054$; visit 4: $t(16) = 3.87$, $p < .001$; visit 5: $t(16) = 4.07$, $p < .001$, visit 6: $t(15) = 9.15$, $p < .001$). In contrast, children with ASD showed comprehension at visit 3, $t(13) = 3.63$, $p = .002$, but not at visit 4, $t(13) =$

0.926, p = .186 or visit 5 $t(11) = 1.32$, p = .107. A trend in the expected direction was observed at visit 6, $t(11) = 1.55$, p = .075.

In sum, TD children displayed evidence of wh-question comprehension by 32 months of age (i.e., visit 4) whereas the ASD group demonstrated significant comprehension at visit 3; however, the ASD group was unable to maintain this level of comprehension consistently for the rest of the visits.

We next examined the number of children in both groups at each visit who demonstrated wh-question comprehension (see Table 5). Difference scores were created for looking during “where” questions minus looking during “what” questions. Positive scores indicated better understanding of wh-questions because these indicate that children looked longer at the matching scene during the ‘where’ questions compared to the ‘what’ questions. Children who showed a difference of 0.40 sec or more between where- and what- trials (in the correct direction) would be designated as “strong” comprehenders, while those who showed a difference of between 0.39 and 0.01 sec were designated as “weak” comprehenders (Goodwin et al., 2012). All children who showed a difference in the wrong direction (i.e. less than zero) were designated “non-comprehenders” (see Table 5). In all the visits for each group, there were more comprehenders than non-comprehenders. We modified Goodwin et al.’s (2012) criterion for “strong” and “weak” comprehenders as participants in our study were at a lower language level compared to his and so participants tended to show smaller differences between the where – and what-trials.

In order to investigate what distinguished comprehenders from non-comprehenders and to reveal individual differences, Pearson’s correlations (see Table 6) were conducted to reveal the extent to which children’s early or concurrent language measures correlated with their later wh-question comprehension scores. Because we had numerous language measures and wh-comprehension scores for subject- and object – questions, we would lose degrees of freedom if we intercorrelated all of these measures. Therefore, in order to increase the power of the study, we used the children's combined (averaged) object-wh and subject-wh scores at each visit, as they performed similarly on these types of questions. Bivariate correlations were performed between these scores and five sets of language measures, including the Vineland, Mullen, CDI, ROWPVT (receptive vocabulary) and EOWPVT (expressive vocabulary); therefore, a Bonferroni correction of $p = .005$ was used. The full set of correlations is presented in Tables 6 and 7. In the TD group, children with higher wh-question comprehension scores at visit 6 had had larger vocabulary scores (CDI) at visits 2 and 3 (rs > .700, ps < .005). Children with greater expressive vocabulary (EOWPVT) at visits 5 and 6 also had higher wh-comprehension scores at visit 6 (rs > .700, $p < .005$). Due to the stricter significance level ($p = .005$) correlations among language measures and wh-question comprehension scores in the ASD group did not reach significance (see Table 7).

Does children's early comprehension of SVO word order predict their later comprehension of wh-questions?

The degree to which children's early understanding of canonical SVO word order predicted later wh-question comprehension was analyzed next. This kind of analysis is potentially perilous because of the small number of participants in each group. For example, any child whose word order data had to be eliminated for a particular visit (for reasons described above) would also have to be eliminated from these regression analyses (N = 8). Therefore, we considered increasing our power by creating a larger dataset which combined the current participants with those of Goodwin et al. (2012; Naigles et al., 2011). The most obvious argument against combining the datasets is that the visual stimuli that our participants saw were different from that which Goodwin et al.'s participants saw; e.g., Goodwin et al.'s participants saw apples and keys hitting flowers and books, while our participants saw horses and birds tickling, pushing, etc. each other. Moreover, the word order stimuli for cohort 2 (Naigles et al., 2011) differed from the word order stimuli for our participants; e.g., the earlier group of children saw a girl and a boy character whereas our participants saw a horse and a bird character. However, the reasons for combining the two sets of wh-question findings are many. First, both their TD children and children with ASD demonstrated the same pattern of understanding that our participants displayed. That is, both their TD children and our TD children displayed stable comprehension of wh-questions by 32 months of age. As for the children with ASD, Goodwin et al.'s group demonstrated comprehension by 54 months of age, and our children were trending towards significance in the same direction by 53 months of age. Second, at visit 6, TD children from their cohort and TD children in our cohort were at the same language level when compared on their Mullen receptive and expressive raw scores (see Table 8). Similarly, the children with ASD from their cohort and our cohort were also at the same language level at their last visit. Third, a similar pattern was observed in both cohorts, such that the language level at which the children with ASD showed comprehension of wh-questions was similar to the language level at which the TD children first showed comprehension of wh-questions. At visit 6, when Goodwin et al.'s (2012) and our participants with ASD first showed comprehension of wh-questions, their language levels were quite similar to those of the TD children, when this group first showed stable comprehension. At visit 3, TD group had a production score of 74% on their CDI, and the ASD group's CDI production score was 66% at visit 6 (Goodwin et al., 2012). As for our participants, there were no significant differences between TD group's receptive and expressive vocabulary scores at visit 4 from ASD group's receptive and expressive vocabulary scores at visit 6 (see Table 10). Fourth, whereas the word order stimuli for the two cohorts were different, the word order layouts themselves were almost identical, with two characters, five common transitive verbs and reversible actions, *push, tickle, wash, hug, and ride*. Fifth, the language levels of the TD children and children with ASD at their first visit were not different for both cohorts (see Table 9). These being deemed sufficient reasons, the combined dataset for the word order-wh-question comparison now included 35 participants in the TD group and 31 in the ASD group.

A bar-plot of each child's score is presented in Figure 4 for the TD group and Figure 5 for the ASD group. The y-axis represents children's proportion of looking to the match during the test trials minus the baseline trials. Both figures reveal considerable within-group variability in word order comprehension, which is a pre-requisite for predicting the variability already observed in wh-question comprehension. That is, if all the participants had shown a robust understanding of SVO word order, then there would be no early variability with which to predict later wh-question comprehension.

Therefore, to investigate whether earlier SVO word order comprehension predicted later wh-question comprehension, we first conducted bivariate correlations between the two measures. In the TD group, one significant correlation (with our stringent Bonferroni correction of $p < .005$) was obtained (see Table 11). That is, early word order comprehension was correlated with subject-wh-questions at visit 5; TD children who performed well on the word order task at the early visits also performed well on subject-wh-questions at this later visit. In the ASD group, one significant correlation was obtained (see Table 12), in which early word order comprehension was positively correlated with object-wh question comprehension at visit 6.

Thus, children with ASD who were better at understanding SVO word order were also the ones who performed better on object-wh questions at this later visit.[2]

We then conducted two stepwise multiple regressions—with each group separately--to assess the degree to which early word order understanding uniquely contributed to later wh-question comprehension. Thus, in the TD group, the model included the children's word order scores, their visit 1 Mullen visual reception and their visit 2 CDI (language) scores; these were used as predictors of the children's visit 5 subject wh-question comprehension scores. The reason for including visual reception is because this measure taps into children's non-verbal IQ, which is an important indicator of the children's ability to attend to and learn from their world. We included visit 2 CDI scores to examine how an early vocabulary measure contributed to their later language processing ability. In the regression model, word order score was the only significant predictor of later subject-question comprehension, $F(1,30) = 4.43$, $p = .044$ (see Table 13). Moreover, children's word order scores accounted for 13% of the variance in their wh-question comprehension, $R^2 = .129$, Adjusted $R^2 = .100$. Overall, early word order comprehension contributed significantly to later wh-question comprehension even when accounting for children's non-verbal IQ and general language at earlier visits.[3]

In the ASD group, the regression model included visit 1 visual reception, visit 2 CDI, and word order score, with visit 6 object-wh-question comprehension score as the dependent variable. The final overall model was significant, $F(2,26) = 4.85$, $p = .016$ (see Table 14). With all measures entered, children's visual reception scores plus word order scores each contributed significantly to the model, jointly accounting for 27% of the variance in wh-question comprehension. Note that children's word order scores received a higher beta weight (standardized coefficient β) compared to their visual reception scores; therefore, children's understanding of word order, plus their cognitive abilities, both played strong roles in their understanding of object-what-questions.

Discussion

In this study, we addressed two main questions: (a) did children with ASD comprehend subject- and object-wh-questions at the same visit or language level as the TD children? and (b) did children's comprehension of SVO word order at visits 1 and 2 predict their later comprehension of wh-questions? Using animate characters engaged in familiar actions, we found that TD children demonstrated comprehension of both subject- and object-questions by 32 months of age (i.e. at visit 4). At this visit, TD children were able to make the distinction that "where" questions referred to the named object in the question whereas subject and object "what" questions did not refer to the named object; this might be viewed as a necessary precursor to understanding wh-movement. Children with ASD showed comprehension of object-wh-questions at visit 6, but not subject-wh-questions even by 53 months of age (i.e. at visit 6); thus, these children seemed unable to understand the difference between "where" questions and subject-wh-questions. The language level of the ASD group at visit 6, when they showed comprehension of only object-wh-questions, was quite similar to those of TD children at visit 4, the earliest visit when these children showed stable comprehension of both object-wh and subject-wh-questions. Moreover, TD children with higher vocabulary scores at visits 2, 3, 5 and 6 were the ones with higher wh-question comprehension scores at visit 6. In contrast, none of the earlier or concurrent language measures were significantly correlated with wh-question comprehension scores for children with ASD. Furthermore, TD children's comprehension of SVO word order at early visits predicted their later performance on subject-wh-questions; earlier comprehension of SVO word order in combination with non-verbal IQ predicted later performance on object-wh-questions for children with ASD.

These results replicated findings of Goodwin et al. (2012) for TD children. TD children from both his study and our study demonstrated stable comprehension of wh-questions at visit 4. But, our results partially replicated findings of Goodwin et al. (2012) for children with ASD. Compared to his children with ASD who demonstrated subject- and object-wh-question comprehension by 54 months of age, our children with ASD showed comprehension of only object-wh-questions by 53 months of age and did not show comprehension of subject-wh-question comprehension at their last visit. Taken together, these findings suggest that using familiar verbs and animate characters did not change the effect found by Goodwin et al. (2012). It seems that the delay in wh-question comprehension for children with ASD was not due to the stimuli used in his study because using simpler stimuli did not result in earlier comprehension by children with ASD. Wh-questions seem to be harder for children with ASD; even though children were only required to *look* at the correct answer, children with ASD still demonstrated impairments in their understanding. Also, in contrast with Goodwin et al. (2012) findings where higher wh-question comprehension scores for children with ASD were correlated with larger vocabularies, no such relationship existed for our ASD group. Children with ASD in our study had an overall lower language level than that of children with ASD tested by Goodwin et al (2012). For example, at visit 3, the CDI scores were lower for our children with ASD and had a smaller variance compared to Goodwin et al (2012). It is possible that lesser variance contributed to non-significant correlations.

These findings also do not support a pragmatic interpretation because using the IPL paradigm, a paradigm that helped to minimize the social demands of the task, did not help children with ASD demonstrate earlier comprehension of wh-questions. Specifically, children with ASD showed inconsistent performance compared to their TD peers. Even by their last visit, children with ASD did not demonstrate understanding of subject-wh-questions, but only object-wh-questions. This shows impairments in their wh-question comprehension.

Our findings supported the grammatical deficit argument of wh-question comprehension as follows: performance on the earlier word order task strongly predicted performance on wh-question comprehension for both TD children and children with ASD. Overall, children's competence on their understanding of the canonical English SVO word order helped them become more efficient in other linguistic processing, specifically the processing of wh-questions. Children who made greater shifts towards the matching screen in the word order task, were the ones who showed larger shifts in looking at the named item from 'where' question to the 'what' question. Children's representation of these sentence forms helped them understand that the wh-movement in a subject-wh or object-wh-question maps onto to either the agent or patient of the action. These findings are evidence that both young TD children and children with ASD use early-developing syntactic knowledge to process the thematic role of wh-words. Therefore, children whose grammatical competence was impaired on this word order task also showed impairments on their wh-question comprehension abilities. These findings are consistent with a number of experimental studies that have shown that both TD children and children with ASD have successfully abstracted English word order sentence frames and applied it to other grammatical constructions with novel verbs (Gertner, Fisher, & Eisengart, 2006; Naigles et al., 2011). Naigles et al. (2011) showed that children with ASD have knowledge of at least one or two abstract sentence frames, as children were able to map novel verbs in transitive sentences to a causative sentence frame, indicating a generalization of pattern to new instances. Also, their speed and efficiency of processing SVO sentences (word order task) eight months earlier predicted performance of children with ASD on this syntactic bootstrapping task, their later ability to use SVO frames to derive meanings about novel verbs, showing that early grammatical ability is predictive of later syntax. This finding strongly supports the grammatical deficit argument for wh-questions. In our study, performance on word order and their non-verbal IQ predicted better performance on object-wh-questions, however; word order was the first best predictor. This implies that the ASD group is not perseverating on one specific word order and can be flexible in switching between word orders. In fact, some children with ASD from Naigles et al. (2011) who showed understanding of at least one or two abstract sentence frames by demonstrating syntactic bootstrapping were the same participants in our study. Thus, it is possible that children are using these sentence frame representations to also process grammatical constructions like object-wh-questions. It may be the case that when children with ASD heard, "The horse is tickling the bird", they noticed that the visual scene consisted of an agent-verb-patient sentence frame. It is possible that these earlier abstracted sentence frames guided children with ASD to look longer at the named item during the object-wh-questions because of the verbs used in the wh-questions. For example, in the object-wh-question, "What is the horse tickling ___?" children with ASD might realize that verbs like 'tickle' involve causation and that it includes a '*tickler*' and a direct object, a '*ticklee*'. Thus, keeping this direct object in mind, children with ASD were able to map the wh-word back to its object referent in the wh-question. Their non-verbal IQ was also the second best predictor of wh-question comprehension in the ASD group. Visual reception (i.e. non-verbal IQ) measures the child's ability to process information. The tasks in the visual reception domain consist of their ability to attend to a picture, to match objects with or without naming, their memory for objects, and their spatial reasoning which are all important factors that can help children integrate the visual and spatial information in the IPL paradigm. Thus, the ability to remember visual sequences and understanding visual information/concepts are useful for sentence comprehension, especially when those sentences are presented using a *looking* paradigm. This finding is also consistent with those of Weismer, Lord and Esler (2010), who showed that nonverbal cognition is a robust predictor of later language abilities in children with ASD. Prior studies have also shown that non-verbal IQ improves due to behavioral intervention in children with ASD (Peters-Scheffer, Didden, Korzilius, & Sturmey, 2011) and this is also applicable to our children with ASD as they were receiving ABA therapy which might have also helped them make gains in their non-verbal cognition.

Our findings revealed that the deficit in wh-question comprehension extended to both types of wh-questions, as children with ASD were delayed and inconsistent in their understanding of both types of wh-questions compared to their TD peers. Children with ASD showed comprehension of only object-wh-questions at visit 6 but not subject-wh-questions. One possibility is that the "blocked" presentation of the subject-wh and object-wh trials made the object-questions easier to understand as these were presented before the subject-questions (Goodwin et al., 2012). In our IPL paradigm, children were first asked all four object-questions together in a block, followed by all four subject-questions in a block. Children with ASD have been shown to have problems with their executive function, which may have influenced their ability to switch sentence frames for the second block of trials (Pennington & Ozonoff, 1996). This could explain their better performance on object-wh-questions because they always appeared first in the video, followed by subject-wh-questions. Moreover, easier comprehension of object-question compared to subject-questions could be due to the simplicity of our task.

Children were required to only *look* at the matching image to show comprehension of the wh-question. They were not required to answer a question (Su et al., 2014) or point to the matching image (Friedmann & Novogrodsky, 2011) which may have minimized social pressure. Furthermore, the stimuli used for the IPL task consisted of two easily distinguishable characters, i.e., a horse and a bird. In prior studies, most of the tasks consisted of two characters of the same type, and one of a different type. This can be deemed to be confusing for atypical populations. Therefore, the simplicity of our task could have fed into children's better performance on object-wh-questions. Stromswold (1995) has also shown that TD children produce object-wh-questions earlier in development than subject-wh-questions. In contrast, children might have demonstrated difficulties with subject-questions since we asked a *what* wh-word to refer to an animate object in the subject-position rather than a *who* wh-word. This could be potentially confusing for children with ASD who are already seeing a non-prototypical action with animate character acting upon another animate character (Slobin, 1982). However, one argument against such an interpretation is that children should also demonstrate difficulties with object-questions as these questions also used a *what* wh-word to refer to an animate object instead of *who,* but we see the opposite results. Thus, these findings are open to alternative interpretations. It would be interesting to examine children with ASD using the same IPL paradigm but with different wh-questions (e.g., *who* and *which*) to further demonstrate whether the grammatical impairments are prevalent among other wh-questions, or whether the prior results were mostly due to stimuli and task demands.

Limitations and Future Directions

There are some limitations to the study. First, it is possible that we made the task harder for children with ASD by using two animate characters engaged in causative actions. As has been shown in prior research, a prototypical action is an animate object performing an action on an inanimate object (Slobin, 1982). Perhaps this also influenced their later comprehension of wh-questions. Second, we are limited in our argument to further distinguish syntactic challenges from pragmatic challenges as this study did not analyze their production of wh-questions or their joint attention skills. Joint attention would be a key predictor of wh-question production to investigate in future studies. Joint attention taps into pragmatic skills in children and therefore it would be important to examine whether joint attention skills are related to later syntactic development. Perhaps, if their joint attention is impaired, then we might also see pragmatic aspects of their wh-question production being impaired. Third, we are restricted in the generalizability of these findings with children with ASD as these children were receiving ABA as their primary intervention, and therefore the generalizability of these findings to the ASD population as a whole are limited. Fourth, it might be possible that we had significant correlations between word order and wh-questions because these were both IPL tasks and therefore, there would be some correlations. It would be important to incorporate another IPL task (either noun bias or syntactic bootstrapping), in order to examine whether these predictive relations would still hold after partialling out the effect of another IPL task. Fifth, the current study only investigated "what" and "where" questions; so we are unclear whether children with ASD would also show deficits in other types of wh-questions. Therefore, these questions need to be investigated in more detail, in terms of comprehension.

In future work, it would be interesting to examine the impairment in wh-questions in other languages, and investigate whether the deficits in understanding such wh-questions also hold for languages that do not require wh-movement to the beginning of the sentence. The Goodwin et al (2012) stimulus materials have been expanded and used with South Korean children with ASD. This is an important step towards making cross-linguistic comparisons with wh-words that remain *in-situ*. Exploring such languages will also shed light on the asymmetry between comprehension of subject-questions and object-questions, since there is no overt wh-movement in Korean language. Therefore, children might show similar performance on both the questions.

In conclusion, using the IPL paradigm helped demonstrate early comprehension of wh-questions in TD children. Children with ASD demonstrated an inconsistent and delayed understanding of wh-questions. The findings of the study lends support to the ‘grammatical’ deficit viewpoint as changing the stimuli and using familiar verbs did not help children with ASD demonstrate earlier comprehension compared to previous results (Goodwin et al., 2012). The results suggest that wh-questions present linguistic challenges to children with ASD that go beyond issues of stimuli. Moreover, their performance on an early grammatical competence task was strongly associated with their performance on wh-question comprehension, indicating that their early grammatical abilities are predictive of their later grammatical competence. Therefore, children who did not perform well on the word order task may not have abstracted the SVO sentence frame, thereby, also raising challenges for wh-question comprehension. Thus, the current study shows that wh-questions seem to be a grammatical deficit in children with ASD.

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Avrutin, S. (2000). Comprehension of discourse-linked and non-discourse-linked questions by children and Broca's aphasics. In: Grodzinsky, Y., Shapiro, P.L., Swinney, D. (Eds.), Language and the Brain. Academic Press, New York, pp. 295-313.

Bloom, L., Merkin, S., & Wooten, J. (1982). Wh-questions: Linguistic factors that contribute to the sequence of acquisition. *Child Development, 53*, 1084-1092.

Brandone, A., Golinkoff, R., Pulverman, R., Maguire, M., Hirsh-Pasek, K., & Pruden, S. (2007). Speaking for the wordless: Methods for studying the foundations of cognitive linguistics in infants. In M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, & M. Spivey (Eds.), Methods in cognitive linguistics (pp. 346-366). Amsterdam: John Benjamins.

Martin, N.A., & Brownell, R. (2000). *Expressive One-Word Picture Vocabulary Test, Third Edition (EOWPVT).* Novato, CA: Academic Therapy Publications.

Martin, N.A., & Brownell, R. (2000). *Receptive One-Word Picture Vocabulary Test, Second Edition (ROWPVT)*. Novato, CA: Academic Therapy Publications.

Candan, A., Kuntay, A.C., Yeh, Y., Cheung, H., Wagner, L., & Naigles, L.R. (2012). Language and age effects in children's processing of word order. *Cognitive Development, 27*, 205-221.

Charman T, Drew A, Baird C, Baird G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). *Journal of Child Language, 30*, 213-236.

Davis, E.A. (1932). The form and function of children's questions. *Child Development, 3*(1), 57-74.

Deevy, P., Leonard, L.B. (2004). Comprehension of wh-questions in children with specific language impairment, Journal of Speech, Language, and Hearing Research, 47, 802-815.

Ebbels, S., & van der Lely, H. (2001). Meta-syntactic therapy using visual coding for children with severe persistent SLI, *International Journal of Language Communication Disorders, 36*, 345-350.

Edelson, L.R., Fine, A, & Tager-Flusberg, H. (2008). Comprehension of nouns and verbs in toddlers with autism: An eye-tracking study. International Meeting For Autism Research, London.

Eigsti, I.M., & Bennetto, L. (2009). Grammaticality judgments autism: Deviance or delay. *Journal of Child Language, 19,* 1-23.

Eigsti, I.M., Bennetto, L., & Dadlani, M.B. (2007). Beyond pragmatics: Morphosyntactic development in autism: *Journal of Autism and Developmental Disorders, 37*, 1007-1023.

Fenson, L., Dale, P., Reznick, J., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. Monographs of the Society for Research in Child Development, 59, 14-24.

Fernald, A., Perfors, A., & Marchman, V.A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year, *Developmental Psychology, 42*(1), 98-116.

Friedmann, N., & Novogrodsky, R. (2011).Which questions are the most difficult to understand? The comprehension of wh-questions in three subtypes of SLI, *Lingua, 121*, 367-382.

Friedmann, N., & Szterman, R. (2011). The comprehension and production of wh-questions in deaf and hard-of-hearing children, Journal of Deaf Studies *and Deaf Education, 16*(2), 212-235.

Friedmann, N., Szterman, R., Haddad-Hanna, M. (2010). The comprehension of relative clauses and wh-questions in Hebrew and Palestinian Arabic hearing impairment, In: Costoa, J., Castro, A., Lobo, M., Pratas, F. (Eds.), Language Acquisition and Development: Generative Approaches to Language Acquisition. Cambridge Scholars Press/CSP, Cambridge, UK.

Golinkoff, R., Hirsh-Pasek, K., Cauley, K., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm: *Journal of Child Language, 14*, 23-45.

Golinkoff, R.M., Ma, W., Song, L., & Hirsh-Pasek, K. (2013). Twenty-five years using the intermodal preferential looking paradigm to study language acquisition: What have we learned? *Perspectives on Psychological Science, 8*(3), 316-339.

Goodwin, A., Fein, D., & Naigles, L. (2012). Comprehension of *wh*-questions precedes their production in typical development and autism spectrum disorders*, Autism Research, 5,* 109-123.

Goodwin, A., Fein, D., & Naigles, L. (2015). The role of maternal input in the development of wh-question comprehension in autism and typical development. *Journal of Child Language, 42*, 32-63.

Hickok, G., & Avrutin, S. (1996). Comprehension of wh-questions in two Broca's aphasics, *Brain and Language, 52*, 314-327.

James, S.L., & Seebach, M.A. (1982). The pragmatic function of children's questions, *Journal of Speech, Language, and Hearing Research, 25*, 2-11.

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

Kukona, A., & Tabor, W. (2011). Impulse processing: A dynamical systems model of incremental eye movements in the visual world paradigm. *Cognitive Science, 35*, 1009-1051.

Lord, C., Rutter, M., DiLavore, P., Risi, S. (1999). The autism diagnostic observation schedule-generic. Los Angeles, CA: Western Psychological Service.

Lovaas, I.O. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*, 3-9.

Maratsos, M. (1998). The acquisition of grammar. In D. Kuhn & R.S. Siegler (Eds.), Handbook of child psychology, Vol. 2: Cognition, perception, and language (pp. 421–466). New York: John Wiley & Sons.

Marchman, V.A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood, *Developmental Science, 11*(3), 9-16.

Mitchell, S., Brian, J., Zwaigenbaum, L., Roberts, W., Szatmari, P., Smith, I., & Bryson, S. (2006). Early language and communication development of infants later diagnosed with autism spectrum disorder, *Developmental and Behavioral Pediatrics, 27*(2), S69-S78.

Mullen, E.M. (1995). Mullen scales of infant development. Circle Pines, MN: American Guidance Service.

Naigles, L., Kelty, E., Jaffery, R., & Fein, D. (2011) Abstractness and continuity in the syntactic development of young children with autism. *Autism Research, 4*, 422-437.

Naigles, L.R., & Tovar, A.T. (2012). Portable intermodal preferential looking (IPL): Investigating language comprehension in typically developing toddlers and young children with autism. *Journal of Visual Experience, 70*, 4331.

Pennington, B.F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 37*(1), 51-87.

Peters-Scheffer, N., Didden, R., Korzilius, H., & Sturmey, P. (2011). A meta-analytic study on the effectiveness of comprehensive APA-based early intervention programs for children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 5*, 60-69.

Piotroski, J., & Naigles, L. (2011). Preferential looking. In E. Hoff (Ed.), Research Methods in Child Language: A Practical Guide (pp. 17-28). Hoboken, NJ: Wiley-Blackwell.

Roeper, T. (2004). Diagnosing language variations: underlying principles for syntactic assessment, *Seminars in Speech and Language, 25*, 41-55.

Rutter, M. (1978). Diagnosis and definition: In M. Rutter & E. Schopler (Eds.), Autism: A reappraisal of concepts and treatment (pp. 1-26). New York: Plenum Press.

Scarborough, H.S., Rescorla, L., Tager-Flusberg, H., Fowler, A.E., Sudhalter, V. (1991). The relation of utterance length to grammatical complexity in normal and language disordered groups, *Applied Psycholinguistics, 12*, 23-45.

Seidl, A., Hollich, G., & Jusczyk, P.W. (2003). Early understanding of subject and object wh-questions. *Infancy, 4*, 423-436.

Slobin, Dan. I. 1982. Universal and particular in the acquisition of language. In Lila. R. Gleitman & Eric. Wanner (eds.), Language Acquisition: The State of the Art. Cambridge, UK: Cambridge University Press.

Sparrow, S.S., Cicchetti, D.V., & Balla, D.A. (2005). Vineland adaptive behavior scales: Second edition (Vineland II), survey interview form/caregiver rating form. Livonia, MN: Pearson Assessments.

Stavrakaki, S. (2006). Developmental perspective on specific language impairment: Evidence from the production of wh-questions by greek SLI children over time, *Advances in Speech Language Pathology, 8*(4), 384-396.

Stromswold, K. (1995). The acquisition of subject and non-subject wh-questions. *Language Acquisition*, 4, 5-48.

Su, Y., Jin, Y., Wan, G.B., Zhang, J.S., & Su, L.Y. (2014). Interpretation of wh-words in Mandarin-speaking high-functioning children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 8*, 1364-1372.

Sussman, R.S., & Sedivy, J.C. (2003). The time-course of processing syntactic dependencies: Evidence from eye movements. *Language and Cognitive Processes, 18*, 143-163.

Swensen, L., Kelley, E., Fein, D., & Naigles, L. (2007). Processes of language acquisition in children with autism: Evidence from preferential looking. *Child Development, 78*, 542-557.

Swingley, D. (2011). Looking while listening. In E. Hoff (Ed.), Research Methods in Child Language: A Practical Guide (pp. 29-42). Hoboken, NJ: Wiley-Blackwell.

Tager-Flusberg, H. (1994). Dissociations in form and function in the acquisition of language by autistic children. In H. Tager – Flusberg (Ed.), Constraints on language acquisition: Studies of atypical children (pp. 175-194). Hillside, NJ: Erlbaum.

Tek, S., Jaffrey, G., Fein, D., & Naigles, L.R. (2008). Do children with autism show a shape bias in word learning? *Autism Research, 1*, 202-215.

Tyack, D., & Ingram, D. (1977). Children's production and comprehension of questions. *Journal of Child Language, 4*(2), 211-224.

van der Meer, M., van Atteveldt, W., Coopmans, P., & Philip, W. (2001). Subject-object asymmetry in child comprehension of wie-questions. *Linguistics in the Netherlands, 18*(1), 167-176.

Venker, C.E., Eernisse, E.R., Saffran, J.R., & Weismer, S.E. (2013). Individual differences in the real-time comprehension of children with ASD. *Autism Research*, *6*, 417-432.

Weismer, S.E., Lord, C., Esler, A. (2010). Early language patterns of toddlers on the autism spectrum compared to toddlers with developmental delay. *Journal of Autism Developmental Disorder, 40*, 1259-1273.

Wetherby, A.M., & Prutting, C.A. (1984). Profiles of communicative and cognitive-social abilities in autistic children. *Journal of Speech and Hearing Research, 27*, 364-377.

Wong, A.M.-Y., Leonard, L.B., Fletcher, P., Stokes, S.F. (2004). Questions without movement: A study of cantonese – speaking children with and without specific language impairment, *Journal of Speech, Language, and Hearing Research, 47*, 1440-1453.

Footnotes

[1] In a more stringent test of object-wh and subject-wh-question comprehension, we compared children's mean percent looking time to the match for the 'control' (trial 2; Table 3) vs. 'what' trials (trials 8 & 34; Table 3). The idea is that children who understand wh-questions should look longer at the matching scene when a what-question is asked compared to when no question is asked. Since we are comparing looking at the bird or the horse during the 'control' trial with the 'what' trials, we analyzed both the object-what and subject-what questions separately, depending on whether the correct target was the bird vs. the horse. The TD group looked significantly longer at the bird during the object-wh-question at visit 3 than during the control trial, $t(16) = 2.20$, $p = .044$, but then demonstrated no preferences at visits 4 – 6, $ts(16) > -.04$, $ps > .05$. No other comparisons were significant for either object-what questions or subject-what questions involving bird and horse in the ASD group (ts(15) < .300, ps > .05). As for children with ASD, these findings relate to their initial preference for the bird at visit 4. For the object-wh and subject-wh questions for which the answer was horse, children did not did not give better responses by looking longer at the horse compared to their baseline trial when they looked longer at the bird.

[2]An additional measure of word order that examined the first look latency to the match during the test trial, i.e., whether SVO latency predicted later wh-question comprehension. However, we found no significant correlations between this latency measure and any wh-question measures.

[3] We also investigated whether TD children and children with ASD's MLUs in production at visit 2 might account for some of the variance in wh-question comprehension. However,

stepwise regressions still chose word order as the most significant predictor of wh-question comprehension amongst children’s non-verbal IQ, CDI, and MLUs.

Table 1. Standardized Test data for Typically Developing (TD) and Autism Spectrum Disorder (ASD) Groups at their First and Final Visits (M, SD)

	TD	ASD	t	p-values
Visit 1				
Gender	13 boys, 4 girls	10 boys, 4 girls		
ADOS	1.47 (1.66)	14.50 (3.70)	-12.21	<.001
Range[a]	0-5	7-21		
CARS	16.21 (1.96)	37.96 (6.10)	-12.81	<.001
Range[b]	15-22.5	31-52		
CDI (Infant version)[c]				
Word Production	123.59 (108.15)	66.21 (113.60)	1.44	.161
Mullen raw scores				
Visual reception	25.88 (3.46)	27.57 (5.37)	-1.06	.299
Fine motor	22.59 (2.60)	25.07 (4.20)	-2.02	.053
Receptive language	22.76 (3.87)	19.64 (10.37)	1.07	.302
Expressive language	20.35 (5.70)	16.29 (6.64)	1.84	.077
Mullen T-Scores				
Visual reception	59.35 (11.37)	36.57 (15.12)	4.79	<.001
Fine motor	53.41 (10.95)	33.43 (16.81)	3.99	<.001
Receptive language	55.53 (13.26)	33.79 (19.62)	3.67	.001
Expressive language	51.71 (15.05)	26.50 (8.86)	5.52	<.001
Vineland standard scores				

Communication	105.12 (9.87)	72.07 (15.45)	7.22	<.001
	TD	ASD	t	p-value
Daily Living	103.76 (9.46)	79.50 (15.05)	5.47	<.001
Socialization	101.71 (6.08)	73.07 (8.53)	10.90	<.001
Motor	98.06 (6.79)	87.64 (14.85)	2.42	.026
Visit 3				
CDI (toddler version)[c]	456.06 (136.69)	178.75 (169.96)	4.79	<.001
Visit 4				
ROWPVT	43.31 (11.96)	32.64 (20.08)	1.80	.083
EOWPVT	115.81 (14.90)	86.36 (24.82)	3.87	.001
Visit 5				
ROWPVT	50.94 (10.60)	40.57 (19.26)	1.80	.087
EOWPVT	120.24 (13.06)	91.79 (23.26)	4.08	.001
Visit 6				
ROWPVT	60.25 (10.70)	48.29 (19.35)	2.06	.053
EOWPVT	125.56 (11.87)	97.07 (23.95)	4.04	.001
Mullen raw scores				
Visual reception	43.56 (4.02)	40.00 (7.67)	1.56	.135
Fine motor	38.56 (5.11)	33.93 (7.11)	2.07*	.048
Receptive language	40.31 (4.88)	34.21 (9.35)	2.19*	.041
Expressive language	39.69 (5.44)	29.57 (13.78)	2.58*	.020

	TD	ASD	T	p-value
Mullen T-scores				
Visual reception	63.81 (11.32)	40.50 (18.97)	4.02*	.001
Fine motor	59.50 (16.32)	31.86 (17.85)	4.43*	<.001
Receptive language	63.13 (10.90)	37.21 (20.27)	4.27*	<.001
Expressive language	59.88 (10.73)	35.00 (22.48)	3.78	.001

*$p < .05$.
[a]Autism spectrum=7+; autism=12+.
[b]CARS range= 15-60; Autism spectrum= 30+; autism= 36+.
[c]Number of words produced out of 396.
ADOS, Autism Diagnostic Observation Schedule; CARS, Childhood Autism Rating Scale; CDI, Communication Development Inventory. ROWPVT, Receptive One-Word Picture Vocabulary Test.

Table 2. Sample Layout of the Word Order Video

Video 1	Audio	Video 2
P[1] Horse waves	Look, a horse! See, the horse!	Blank
P Blank	Look a bird! See, the bird!	Bird waves
P Horse waves	We see both!	Bird waves
P Horse waves	Look at the horse!	Bird waves
P Horse waves	Look at the bird!	Bird waves
1 Blank	Look, pushing! See, pushing!	Bird pushes horse
2 Horse pushes bird	Look, pushing! Wow, pushing!	Blank
3 Horse pushes bird	They are on both screens!	Bird pushes horse
4 Horse pushes bird	Look, the bird is pushing the horse!	Bird pushes horse
	(Block repeats with tickle/pull/wash/ hug/ride)	

[1]*P* indicates the pretest trials

Table 3. Sample Layout of the Wh-Question Video

Trial Type	Audio	Video 1	Video 2
1	Oh, look!	Black	Black
2 Control-Baseline	They're on both screens!	Bird	Horse
4 Familiarization	Look at this!	Horse tickles Bird	Black
6 Familiarization	See this?	Black	Horse tickles bird
8 Test[a]	What did the horse tickle __?	Bird	Horse
9-28	(Block repeats with wash/hug/ride)		
30 Familiarization	Look at this!	Bird hugging horse	Black
32 Familiarization	See this?	Black	Bird hugging horse
34 Test[b]	What hugged __ the horse?	Bird	Horse
35-53	(Block repeats with ride/tickle/wash)		
54	Isn't this fun?		
56 Where-Test[c]	Find the horse!	Bird	Horse
58 Where-Test[c]	Find the bird!	Bird	Horse

✓= Red dot flashing to draw the child's attention back to the center before the next trial begins.
ǂ = Fish swimming across screen to maintain children's interest.
[a]Object-wh-questions = What did the horse tickle?; What did the bird wash?; What did the bird hug?; What did the horse ride?
[b]Subject-wh-questions = What hugged the horse?; What rode the bird?; What tickled the bird?; What washed the horse?
[c]Where is the horse?; Where is the bird?

Table 4. Children's Ages at Each Visit (in Months)

	TD		ASD	
Visit	M	(SD)	M	(SD)
1	19.74	(1.25)	32.93	(7.28)
2	24.54	(1.22)	36.98	(7.63)
3	28.52	(1.39)	41.41	(7.31)
4	32.33	(1.35)	45.71	(7.19)
5	36.51	(1.51)	48.97	(7.67)
6	40.69	(1.53)	53.24	(7.25)

Table 5. Number of Children Showing Strong, Weak, or No Comprehension of Wh-Questions (Subject – and Object – Questions Averaged)

Visit	Comprehension Type	TD	ASD
Visit 3	Strong	0	1
	Weak	**13**	**11**
	None	4	2
Visit 4	Strong	5	2
	Weak	**9**	**6**
	None	3	6
Visit 5	Strong	3	2
	Weak	**12**	**5**
	None	2	7
Visit 6	Strong	4	2
	Weak	**12**	**5**
	None	*0*	*7*

Note. Strong: $x > .40$ difference score; Weak: $.39 > x > .01$ difference score; None: $0 > x$ difference score. ASD, autism spectrum disorder; TD, typically developing.

Table 6. Cross-lagged and Concurrent Pearson Correlations Between Language Measures and Wh-Question Comprehension For TD Children Across All Visits (N=17).

	Visit 3	Visit 4	Visit 5	Visit 6
Variable	Where – Test Combined	Where – Test Combined	Where – Test Combined	Where – Test Combined
Visit 1				
MSEL	.238	.294	.074	.425
VABS	.070	-.358	-.162	.343
CDI	.084	.298	-.271	.623
Visit 2				
VABS	-.242	-.444	.128	.369
CDI	.340	.478	-.148	.714*
Visit 3				
VABS	-.071	-.111	.202	.641+
CDI	.077	.302	-.153	.858*
Visit 4				
VABS		-.456	.145	.460
ROWPVT		.451	.182	.579
EOWPVT		.534	.131	.592
Visit 5				
VABS			.229	.541
ROWPVT			.177	.504

	Visit 3	Visit 4	Visit 5	Visit 6
Variable	Where – Test Combined	Where – Test Combined	Where – Test Combined	Where – Test Combined
EOWPVT			.102	.733*
Visit 6				
MSEL				.554
VABS				.380
ROWPVT				.639+
EOWPVT				.780*

Note. MSEL = Mullen Scales of Early Learning Composite; VABS = Vineland Adaptive Behavior Scales Composite; CDI = Communicative Development Inventories; ROWPVT = Receptive One-Word Picture Vocabulary Test; EOWPVT = Expressive One-Word Picture Vocabulary Test. *p < .005, two-tailed; +*p* < .01.

Table 7. Longitudinal Pearson Correlations Between Language Measures and Wh-Question Comprehension For Children with ASD Across All Visits (N=14).

	Visit 3	Visit 4	Visit 5	Visit 6
Variable	Where – Test Combined	Where – Test Combined	Where – Test Combined	Where – Test Combined
Visit 1				
MSEL	.196	.319	.308	.020
VABS	.315	.372	-.016	.052
CDI	.247	.394	.393	.352
Visit 2				
VABS	.355	.395	.115	.027
CDI	.267	.526+	.409+	.354
Visit 3				
VABS	.283	.365	-.078	-.029
CDI	.083	.302	-.099	.101
Visit 4				
VABS		.411+	.045	.004
ROWPVT		.352	.139	-.044
EOWPVT		.467+	.438+	-.053
Visit 5				
VABS			.147	.020
ROWPVT			.107	-.095
EOWPVT			.239	.089

	Visit 3	Visit 4	Visit 5	Visit 6
Variable	Where – Test Combined	Where – Test Combined	Where – Test Combined	Where – Test Combined
Visit 6				
MSEL				.124
VABS				-.105
ROWPVT				-.125
EOWPVT				.150

Note. MSEL = Mullen Scales of Early Learning Composite; VABS = Vineland Adaptive Behavior Scales Composite; DLS = Daily Living Skills; CDI = Communicative Development Inventories; ROWPVT = Receptive One-Word Picture Vocabulary Test; EOWPVT = Expressive One-Word Picture Vocabulary Test. *p < .005, two-tailed; $^{+}$represents all the correlations greater than $r = 0.40$, $0.05 < ps < .16$.

Table 8. Comparison of TD and ASD Participants from Both Cohorts at Visit 6 on the MSEL.

Visit 6	Goodwin et al(2012)	Current Study	t	p-values
TD	*M (SD)*	*M (SD)*		
Mullen Receptive	38.67 (4.13)	40.31 (4.88)	-1.07	.295
Mullen Expressive	39.72 (5.49)	39.69 (5.44)	.018	.985
ASD				
Mullen Receptive	31.18 (10.78)	34.21 (9.35)	-.828	.414
Mullen Expressive	27.06 (13.31)	29.57 (13.78)	-.515	.611

Table 9. Comparison of TD and ASD Participants from Both Cohorts at Visit 1 on Standardized Tests.

Visit 1	Goodwin et al(2012)	Current Study	t	p-values
TD	*M (SD)*	*M (SD)*		
CDI	118.78 (114.35)	123.59 (108.15)	-.128	.899
Mullen Receptive	25.33 (2.93)	22.76 (3.87)	2.22*	.033
Mullen Expressive	19.44 (4.46)	20.35 (5.70)	-.527	.602
ASD				
CDI	94.12 (111.38)	66.21 (113.60)	.688	.497
Mullen Receptive	23.18 (8.19)	19.64 (10.37)	1.06	.298
Mullen Expressive	18.53 (8.13)	16.29 (6.64)	.829	.414

CDI, Communication Development Inventory; *$p < .05$.

Table 10. Comparison of Language Level between ASD Participants (Visit 4) and TD Participants (Visit 6).

	ROWPVT	EOWPVT	t	p-values
Visit 4				
TD	43.31 (11.96)	31.18 (9.89)	-.859	.398
Visit 6				
ASD	48.29 (19.35)	30.00 (24.56)	.168	.868

Table 11.Cross-lagged Pearson Correlations Between Word Order and Wh-Question Comprehension For TD Children Across All Visits (N=35).

	Visit 3			Visit 4			Visit 5			Visit 6		
Variable	Where-Object	Where-Subject	Where - Test Combined	Where-Object	Where-Subject	Where-Test Combined	Where-Object	Where-Subject	Where-Test Combined	Where-Object	Where-Subject	Where-Test Combined
Word Order	-.291	-.232	-.320	-.147	.151	-.040	.019	.359*	.177	-.057	.262	.083

Note. *p < .05.

Table 12. Cross-lagged Pearson Correlations Between Word Order and Wh-Question Comprehension For Children with ASD Across All Visits (N=31).

	Visit 3			Visit 4			Visit 5			Visit 6		
Variable	Where-Object	Where-Subject	Where - Test Combined	Where-Object	Where-Subject	Where - Test Combined	Where-Object	Where-Subject	Where - Test Combined	Where-Object	Where-Subject	Where - Test Combined
Word Order	.283	.168	.258	-.185	.079	-.102	-.326	-.122	-.268	.381*	.157	.269

Note. *p < .05.

Table 13. Stepwise Regression Analysis for Variables Predicting Visit 5 Subject – What Question Comprehension in TD Children (N=31).

Variable	*B*	*SE(B)*	*β*	*t*	*p*	ΔR^2
Model 1					.044	.129
Word Order	73.12	34.76	.359	2.10	.044	
Excluded Variables						
Visit 1 Visual Reception			-.132	-.714	.481	
Visit 2 CDI			.053	.294	.771	

Table 14. Stepwise Regression Analysis for Variables Predicting Visit 6 Object – What Question Comprehension in Children with ASD.(N=28)

Variable	*B*	*SE(B)*	*β*	*t*	*p*	ΔR^2
Model 1					.041	.145
Word Order	77.39	36.10	.381	2.14	.041	
Model 2					.016	.272
Word Order	103.05	36.06	.508	2.86	.008	
Visit 1 Visual Reception	1.71	.805	.377	2.12	.044	
Excluded Variables						
Model 1						
Visit 2 CDI			.275	1.55	.133	
Visit 1 Visual Reception			.377	2.12	.044	
Model 2						
Visit 2 CDI			.008	.032	.975	

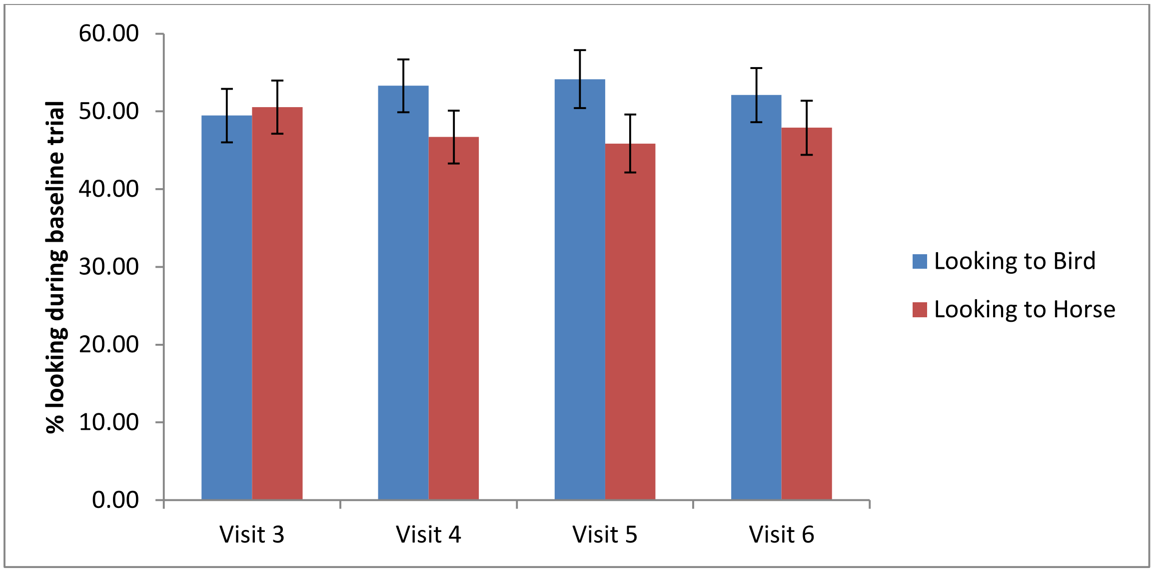

Figure 1a. Percent looking to the bird vs. horse during the control trial for TD children.

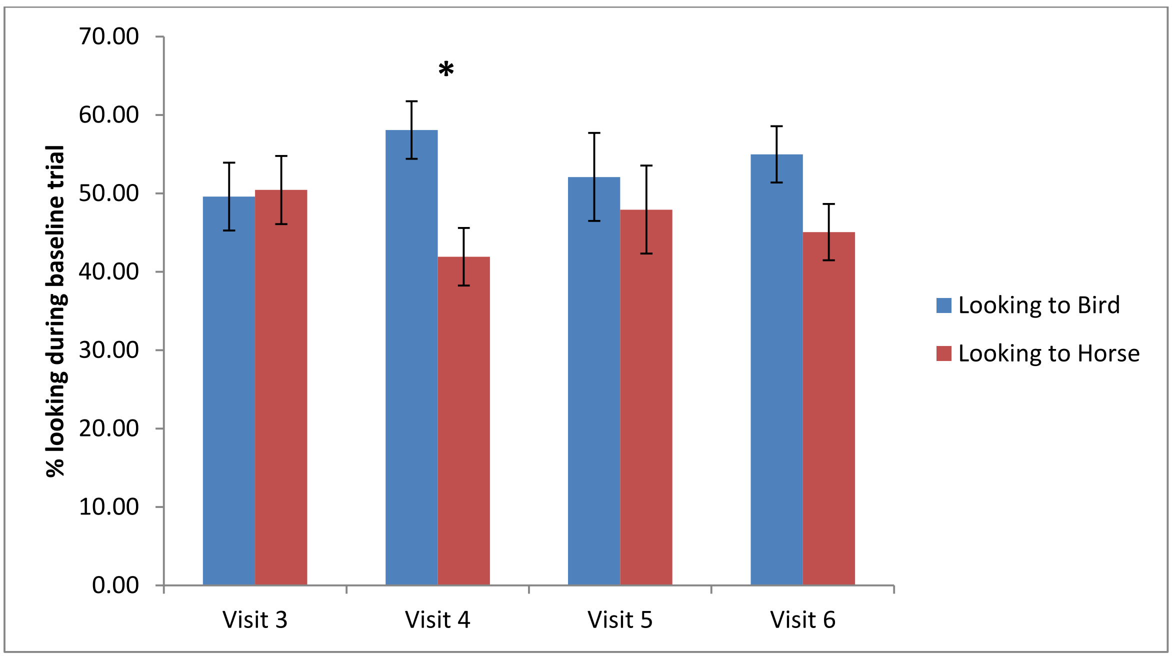

Figure 1b. Percent looking to the bird vs. horse during the control trial for children with ASD, *$p < .05$.

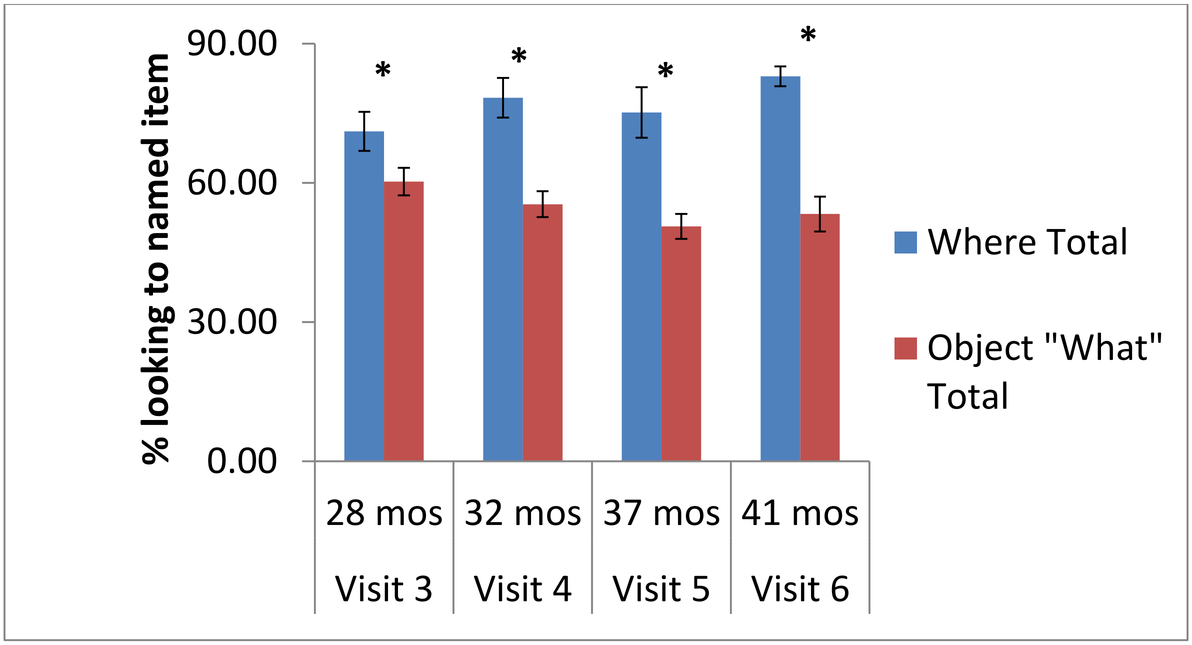

Figure 2a. Comparison of where vs. object-what trials for TD children across visits. *$p < .05$.

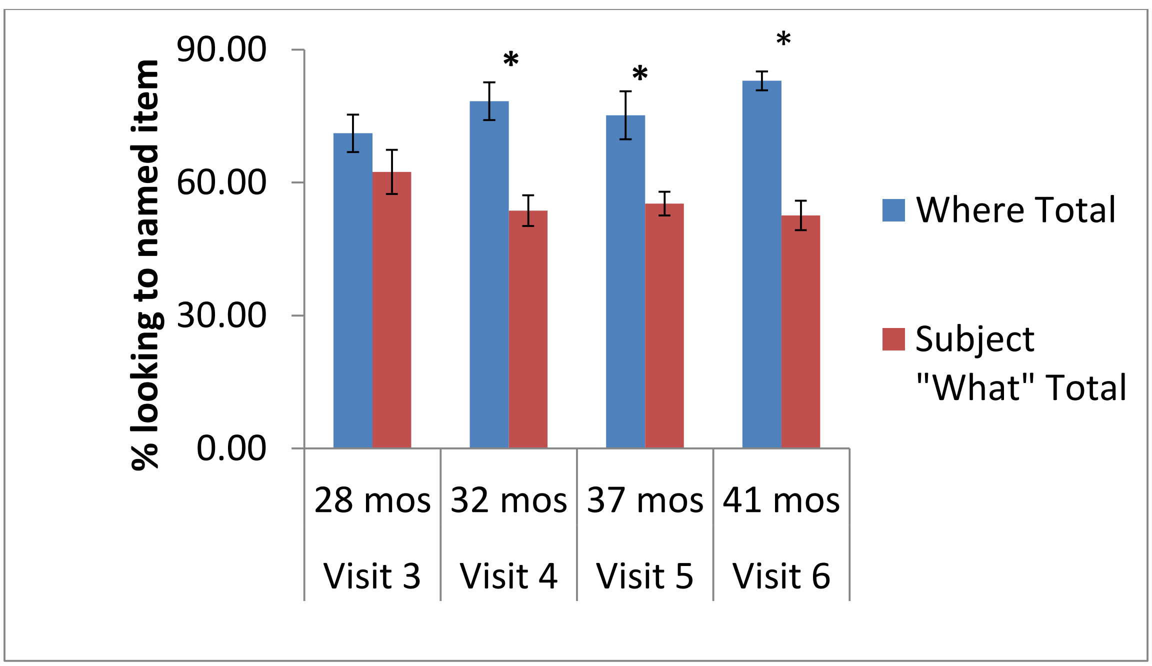

Figure 2b. Comparison of where vs. subject-what trials for TD children across visits. *$p < .05$.

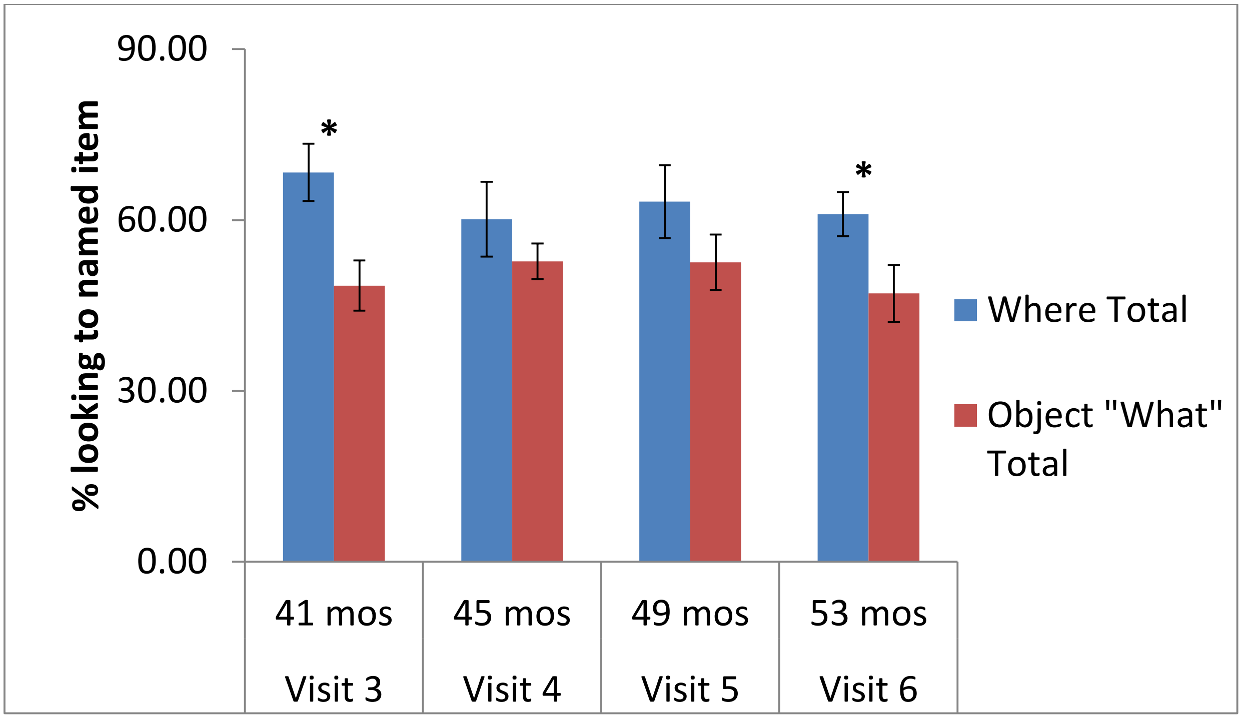

Figure 3a. Comparison of where vs. object-what trials for children with ASD across visits.
*$p < .05$.

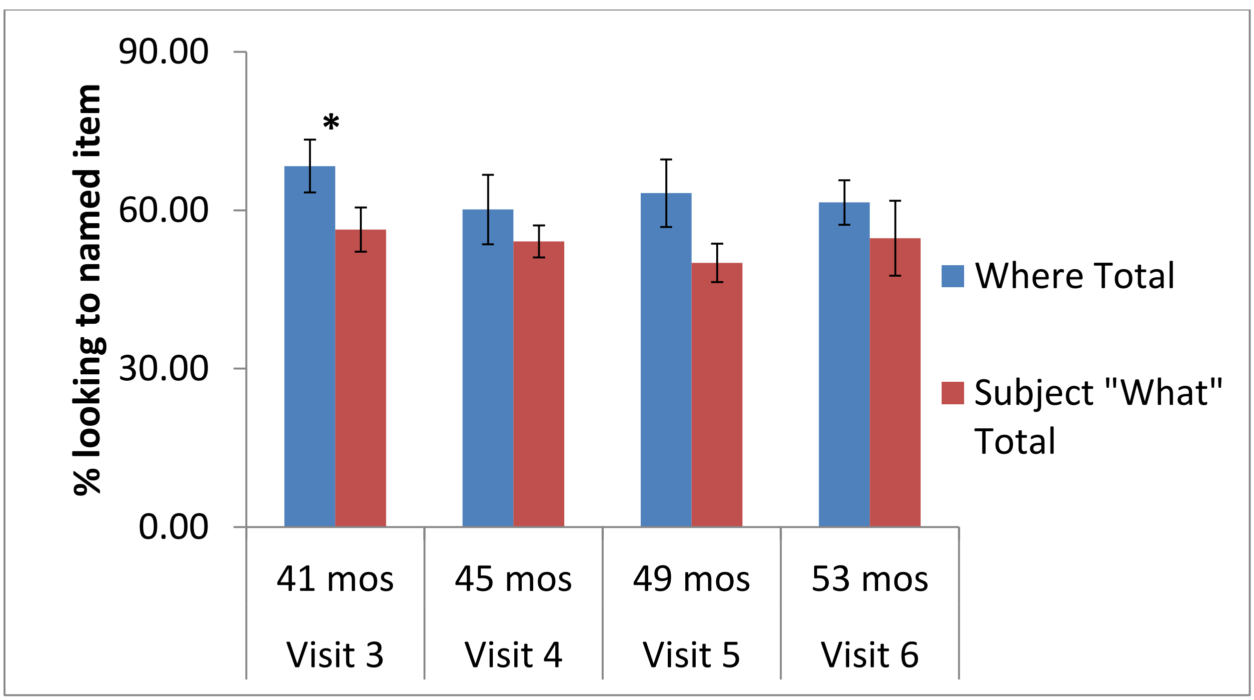

Figure 3b. Comparison of where vs. subject-what trials for children with ASD across visits.
*$p < .05$.

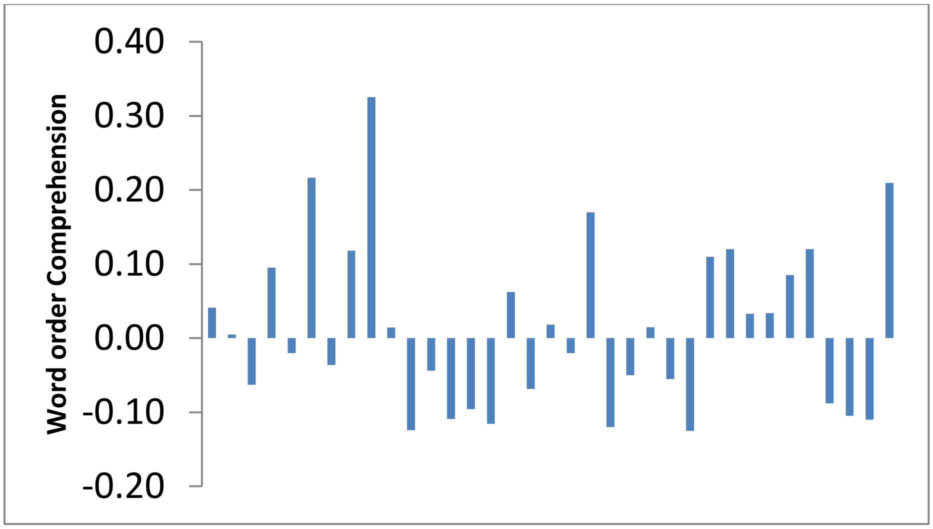

Figure 4. Variability of word order comprehension (Percent looking to match during test minus baseline)in TD children, Visits 1 and 2 combined.

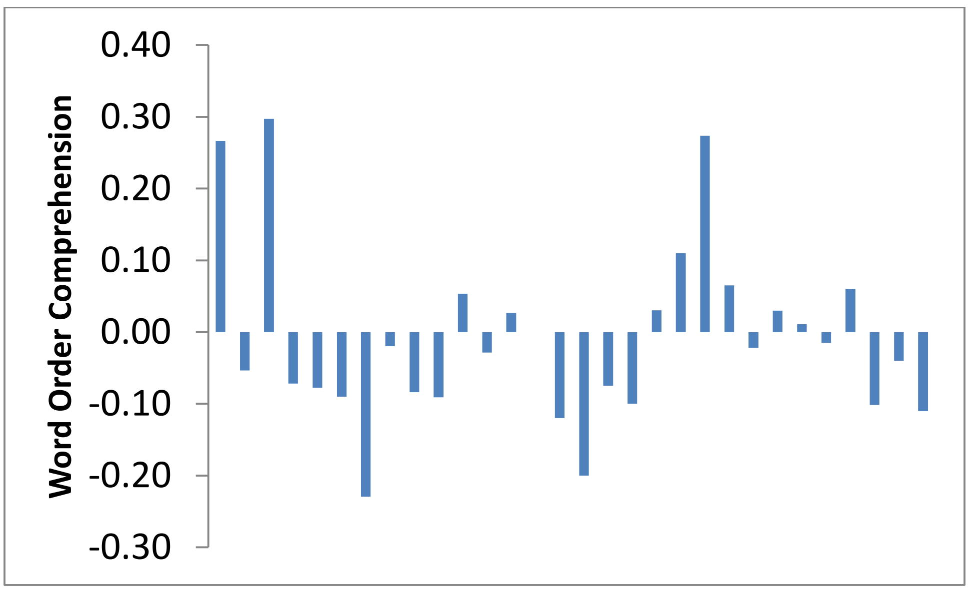

Figure 5. Variability of word order comprehension (Percent looking to match during test minus baseline)in ASD children, Visits 1 and 2 combined.

Appendix A

Percent looking to match for word order video for TD children and children with ASD combined across Visits 1 and 2.