5-12-2022

Measuring referential communication dynamically in older children with ASD

Caitlin Lee
James Madison University

Measuring Referential Communication Dynamically in Older Children with ASD

Caitlin Lee

A thesis submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Master of Science in Speech-Language Pathology

Department of Communication Sciences and Disorders

May 2022

FACULTY COMMITTEE:

Committee Chair: Dr. Emily Zane, Ph.D.

Committee Members/ Readers:

Dr. Geralyn Timler, Ph.D., CCC-SLP

Dr. Jaime Lee, Ph.D., CCC-SLP

Dr. Trevor Stokes, Ph.D.

TABLE OF CONTENTS

List of Tables

List of Figures

Abstract

Research finds individuals with autism spectrum disorder (ASD) are relatively ineffective and/or inefficient at referential communication. However, this research typically uses static metrics of efficacy (how accurately messages were relayed) and efficiency (overall word count), rather than dynamic ones (e.g., Does the speaker alter subsequent descriptions when the listener previously misunderstood them?). The aim of this research is to use dynamic measures of efficacy and efficiency to examine how speakers with and without ASD adjust their message to meet listener needs across time. Fifteen older children with (n = 8) and without (n = 7) ASD were included. Participants interacted with two research assistants (RA1 and RA2). RA1 sat beside the participant. RA2 sat across from them, behind an opaque barrier. A board was positioned before the participant with a doll in its center and four black, ten-sided shapes surrounding the doll. RA1 explained that shapes were configured around RA2 the same way they were the doll. It was the participant's task to use language to guide RA2 to select targeted shapes. There were 64 trials, and each trial was coded by the type of strategy. We created an "efficacy quotient" (EQ), which assigned 1s for trials when participants: 1) changed strategies when the previous trial was unsuccessful, 2) maintained strategies when it was successful. Remaining trials earned a 0. We summed 1s and divided this value by 63 (participants had opportunities to switch/maintain strategies from the second trial on). T-tests were used to compare EQs between groups. We measured the number of words within trials, where shorter utterances in later trials suggest referential shortening. We used linear modeling to compare slopes between groups. We found no significant difference in EQ between groups but found a significant effect of trial, indicating that both groups were equally effective at adjusting communication strategies based on previous success. However, participants with ASD were marginally more efficient: they were quicker to abbreviate utterances to reflect increased listener understanding. Findings emphasize the importance of examining interactions dynamically; such measures capture the realities of turn-taking and may identify strengths in autism that have heretofore gone unnoticed.

INTRODUCTION

Referential Communication

Referential communication refers to the linguistic forms speakers use to enable a listener to identify something (Yule, 1996). For example, the name "Ektorp" is likely unrecognizable to most people, but to an IKEA employee, it is immediately recognized as one of the company's most popular sofas. For speakers to reference clearly to their listener, they must carefully attend to their listener's needs. The sofa owner may say to his friend, "what do you think of my new couch?", if he is unsure about how versed his friend is with IKEA. However, if he knows that his friend also has an Ektorp (or that his friend works at IKEA), the owner might ask, "what do you think of my new Ektorp?" If his friend responds with confusion, the speaker might clarify, "that's the IKEA model name for this couch", to assist in the comprehension breakdown. However, if the speaker receives an appropriate reply, he can continue with, "I made sure to get the newer model with the taller back."

As exemplified above, attuning to a listener's initial needs and subsequent responses is required to effectively communicate reference. The speaker must take into account their partner's status, knowledge, and feelings and use that information to present the content in a way that will be effective for their specific partner (John et al., 2009). Therefore, referential communication is a pragmatic language skill, as it has communicators relying upon the contexts of the situation in order to respond appropriately. This pragmatic element of communication has been thought to depend on Theory of Mind (ToM), the ability to infer the full range of mental states that cause action as well as the ability to reflect on the contents of one's own and another's mind, while more recent research claims that domain general cognition, i.e., executive functioning (EF), underlies referential communication skills (Schuh et al., 2016).

Audience Design & Common Ground

Clark et al. (1982) refers to the way in which speakers tailor their message to accommodate the needs of their specific listener and to accommodate the needs of a specific situation (so that it is successful) as audience design (sometimes “addressee accommodation”, Horton and Keysar [1996]). Krauss and Weinheimer (1964) find that audience design in communicative situations require the speaker to contemporaneously update his or her message to meet the listener’s comprehension. Horton and Keysar support this in a study (1996), where they compared two theoretical models (with one model assuming the speaker adjusts their message to accommodate their listener in both the initial utterance planning and the monitoring stage while the other only assumes the speaker accommodates in the monitoring stage) and found greater evidence with the model stating that speakers plan speech “egocentrically” and then subsequently deliver the plan according to the listener. That is, the speaker, first, plans their message, then modifies it to accommodate the other person. Thus, speakers monitor and adjust the speech plan according to the listener. These adjustments are based on the speaker’s approximate understanding of the listener’s momentary knowledge, beliefs, and thoughts. This skill is also known as common ground.

Common ground refers to the speaker’s presuppositions of what the listener does (and does not) know. It describes the tendency of speakers to modify how they communicate based on the assumption that they share information and knowledge with their listener (De Marchena and Eigsti, 2016). Common ground can be reflected in a number of ways. The speaker can explicitly refer to common ground (for example, from the previous IKEA example, “‘Ektorp’ is the style of couch from IKEA”). Or, a speaker will begin systematically referring to a specific referent with a certain label, once the speaker assumes that the listener understands that label (for example, simply naming “Ektorp” without the descriptor after introducing the couch earlier). However, it could also be implicitly determined; tacitly understood by both the speaker and audience (for example, the speaker is talking to an IKEA employee and names the couch style in conversation). All of these examples are present in many social situations. Once common ground is settled, the current form of the components of common ground becomes the basis on which the speakers’ perspectives are adjusted and updated during language processing (Barr and Keysar, 2006).

Once reference has been successfully established and there is shared knowledge between speaker and listener, referential phrases can shorten (aptly termed "referential shortening"). For example, a speaker may first say, "I just bought an Ektorp, which is a style of Ikea couch;", however, once reference has been established, the underlined phrase can be shortened to "my Ektorp" or "my couch" in subsequent references. Thus, audience design and common ground can be implicitly reflected through referential shortening effect. In two studies by Krauss and Weinheimer (1964, 1966) subjects were required to communicate about novel, nonrepresentational figures. They found a negative relationship between description length and referential frequency, whereby the number of words used to refer to a given referent would decrease with the number of times the speaker/listener communicated about that referent. Initially, speakers in both studies would describe a figure in detail, for example, "the upside-down martini glass in a wire stand," then reduce the description to "inverted martini glass," "martini glass," and, finally, "martini" in successive turns (Krauss and Weinheimer, 1966). This transition from explicit descriptions to shorthand labels is an aspect of audience design that demonstrates shared common ground between speaker and listener.

Importantly, smooth execution of audience design and common ground is dependent upon the listener's comprehension. If the listener were to not understand the speaker's "upside-down martini glass in a wire stand" attempted label, the speaker should not continue with that label, let alone shorten it to "martini glass" in the following trial. Thus, a speaker's message is affected by the confirmation and feedback from the listener. When indications that the speaker's audience is not following the speaker arise, the speaker must switch their referential strategy in order to be effectively communicative. We hereby refer to this as *effective strategy switching*. And, vice versa, when a listener provides confirmation that a message has been successfully received, generally, the speaker should maintain or not deviate far from their referential strategy. We refer to this as *effective strategy maintenance*. Effective strategy switching and maintenance is what directly characterizes effective audience design.

Referential Communication in ASD

As is clear from the previous section, effective referential communication depends on careful attention to the listener's needs, which requires theory of mind and perspective-taking abilities. Further, referential communication requires pragmatic skills, like audience design and common ground. One population who has been shown to struggle with all these skills (pragmatics, perspective-taking, and theory of mind) are individuals with autism spectrum disorder (ASD; Baron-Cohen, 2000; Tager-Flusberg et al., 2005; Landa, 2000; American Psychiatric Association, 2013). Weak or lacking theory of mind skills are cited as the explanation for challenges with perspective-taking for individuals with ASD (Baron-Cohen, 2000). While a systematic review found that children with ASD do not demonstrate significant weakness with perceptual perspective-taking, there was weakness with conceptual perspective-taking (Baron-Cohen, 2000)[1]. Marvin et al. (1976) defines conceptual perspective-taking as taking on the viewpoint of another person's thoughts, feelings and attitudes. This deficit coincides with more egocentric thinking, which may result in breakdowns of establishing common ground and effectively audience designing.

Accordingly, numerous studies have found that individuals with ASD present with referential communication difficulties. In a systematic review, Malkin et al. (2018) summarized and analyzed research findings on verbal referencing abilities of older adolescents and adults with ASD. In this report, 17 of 19 studies on "listener needs/audience design" reported verbal referencing skills that diverged from neurotypical counterparts. Some of these studies find that individuals with ASD specifically struggle to adapt their communication to the needs of their listener, resulting in inefficient communication (Fukumura, 2016; Nadig et al., 2015; Nadig et al., 2009). For example, Fukumura (2016) assessed the sensitivity for referential ambiguity of children and adolescents with ASD while communicating in a shared context and privileged context. In the shared context,

[1] Some of the research from the Baron-Cohen (2000) review does not find similar deficits in perceptual perspective-taking, which is defined by Marvin et al. (1976) as understanding the perceptually visual or auditory experience of another person.

competing pictures (a large door and a small door) and two other images were visible to both the speaker and addressee, but in the privileged context, while the speaker saw both doors, only the "small door" was visible to the addressee. Therefore, the speaker should use contrastive adjectives in the shared context and should not in the privileged perspective. While it was found that ASD speakers used contrastive adjectives (e.g., "big" or "small") to disambiguate between pictures as often as their NT peers in the shared context, they used these adjectives significantly more often than NT speakers in the privileged context. In this case, continuing to use contrastive adjectives did not necessarily break down the speaker's message, but communicating in this overly informative way suggests that ASD speakers were aware of the needs from their perspective and unaware of the changed needs of the listener. This is further supported in studies by Volden et al. (1997), Dahlgren and Dahlgren Sandberg (2008), and Nadig et al. (2009) that individuals with ASD provided less *informative* descriptions of stimuli, less context-appropriate descriptions, and a greater proportion of *irrelevant* features or redundant responses as compared to NT participants when describing visual objects[2]. With a different measure, De Marchena and Eigsti (2016) had similar conclusions about ASD referential communication. They compared the amount of referential shortening (indicative of common ground, as stated earlier) between conditions when the speaker and listener shared background knowledge (and therefore speech can be abbreviated) vs. when they did not (and therefore speech should be longer to provide background information). They found that the ASD group showed less of a difference in length between conditions.

In summary, previous research on referential communication in individuals with ASD has found significant weaknesses in efficient and effective communication as their messages are often characterized as overly descriptive, consisting of non-discriminatory details, and equal in length regardless of the listener's familiarity of the topic. However, the research has consisted of groups that were not well matched or were only matched by age and has been largely concerned with

---

[2] This was found to be true even for the participants who had intact perspective-taking abilities, allowing for successful referential communication in the Volden et al. (1997) study.

individual responses rather than analyzing *how* the responses of the speakers with ASD change throughout discourse. To date, no research has explicitly investigated their underlying skills (strategy monitoring and adjustment) to dynamically measure their success in a referential task. Our study aimed to examine referential communication abilities between ASD and NT children that are well matched in age, language, and intelligence, using new measures (*strategy switching* and *strategy maintenance*) to capture an individual's awareness. Further, we modified how we measured referential shortening to see utterance shortening progression over time, trial by trial (rather than comparing utterance length in condition A vs. condition B).

Current Study

The current study explores referential communication skills in children with and without ASD during a traditional barrier task. Although the task involved a reference board for participants to refer to and provide directional clues, they were not explicitly instructed to use spatial terms. Using spatial-language (with an accurate perspective) would have been the most efficient method of completing this task, however, several children in both groups attempted to describe the physical characteristics of the stimuli to the experiment partner/research assistant. Since the participants who used shape descriptions had ample opportunity to switch their strategy to a more efficient one, these object description trials allowed us to test how participants responded, or adapted, to negative feedback (i.e., partner chose the incorrect shape). In other words, their method of completing the task allowed us to analyze how both ASD and NT children strategy switched/maintained.

Our research questions were as follows: (a) Do children with ASD show similar rates of effective strategy switching and maintenance as compared to NT peers? (b) Do children with ASD show referential shortening effects to the same degree as their NT peers, reflecting an understanding of increased shared knowledge as the experiment progressed? Three hypotheses were made for this subset of participants. First, we predicted that children with ASD would show lower audience design skills (as measured by effective switching and maintenance) than NT peers. Second, we predicted that participants with ASD would show specific challenges with strategy switching (as compared both to their own levels of effective maintenance and as compared to NT children), since it involves EF skills (more so than strategy maintenance), such as working memory and cognitive flexibility, which has been shown to pose as a challenge for children with ASD (Corbett et al., 2009). Third, based on previous findings, we predicted that children with ASD would show relatively lower levels of referential shortening effects (De Marchena and Eigsti, 2016). In order to test our predictions, we compared participants' effective strategy maintenance or switching (depending on the communication partner's feedback) and referential shortening effects (measured by changes in number of words used to target a shape).

METHODOLOGY

Participants

Participants were 8 adolescents with autism spectrum disorder (ASD) and 7 neurotypical (NT) adolescents between the ages of 10.1 and 17.2 years (mean = 13.2 years). They were recruited through Facebook, parenting group blogs, and Craigslist. Prior to coming into the lab, NT participants were screened through the phone to ensure they did not have any of the following comorbid diagnoses: neurological conditions, language impairment, cognitive impairment, or attention-deficit/hyperactivity disorder (ADHD). NT participants who had siblings with ASD were also excluded. Among the ASD group, ASD diagnoses were confirmed in the lab using the Autism Diagnostic Observation Schedule (ADOS-2: Lord et al. 2012) by administrators in the lab who achieved research reliability with a certified trainer. Participants included in the study had to be in mainstream classrooms, have a normal or higher-than-normal IQ, and have normal or higher-than-normal language. Basic language and cognitive abilities were also assessed using the Core Language Subtests of the Clinical Evaluation of Language Fundamentals, 5th Edition (CELF-5: Wiig et al. 2013) and the Kaufman Brief Intelligence Test, 2nd Edition (KBIT-2: Kaufman and Kaufman 2004). Participants whose CELF-5 or KBIT-2 scores were -1.5 standard deviations below the mean were excluded. The ASD and NT groups did not differ significantly on IQ or standardized language skills, nor did they differ in age or gender (Table 1) Finally, all participants' social communication skills were assessed using the Social Communication Questionnaire - Lifetime (SCQ – Lifetime: Rutter et al. 2003).

Table 1: Participant Demographics

	ASD; *n = 8*	**NT; *n = 7***	***p***
Age (months)	155.5 [121-206]	161.71 [130-195]	.65
Male : Female	7:1	6:1	1
KBIT	115.13 [84-129]	115 [96-138]	.99
CELF	112.5 [90-133]	115.57 [85-143]	.76
SCQ - Lifetime	16.63 [12-27]	2.57 [0-7]	< 0.05

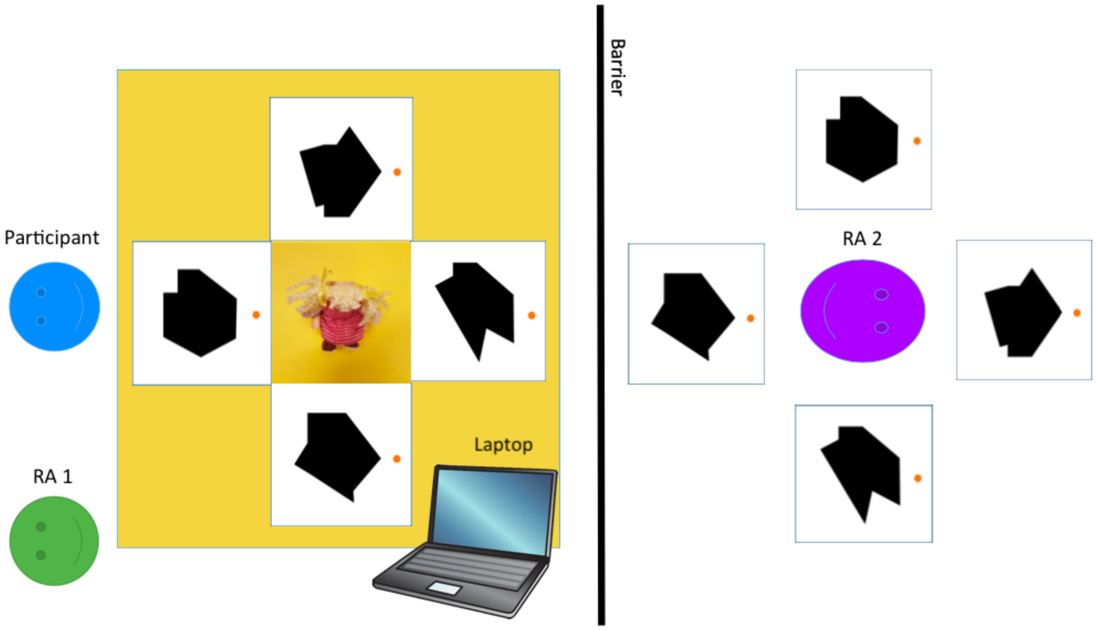
Note. Data presented as mean [range]; KBIT = Kaufman Brief Intelligence Test, 2nd Edition; CELF-5 = Core Language Subtests of the Clinical Evaluation of Language Fundamentals, 5th Edition; SCQ - Lifetime = Social Communication Questionnaire – Lifetime.

Procedures

Set-up

There were three people involved in this task: The participant and two RAs (RA1 and RA2). RA1 and the participant sat down on the floor together in front of a 5-foot partition. RA2 sat on the other side of the participant so that she was not visible to the participant and vice versa (Figure 1). Once the three people were positioned on their respective sides of the partition, RA1 placed a reference board in front of the participant, with a doll sitting in the center of the board. The doll was surrounded by four black, ten-sided shapes, all of which had an orange dot on one side to indicate which side was the "top" of the shape. On the other side of the barrier, the same four shapes were placed around RA2 in same configuration as they were around the doll. For example, in Figure 1, the shape in front of the doll is the same shape as the one in front of RA2. In front of the participant and RA1 there was a laptop sitting on the floor.

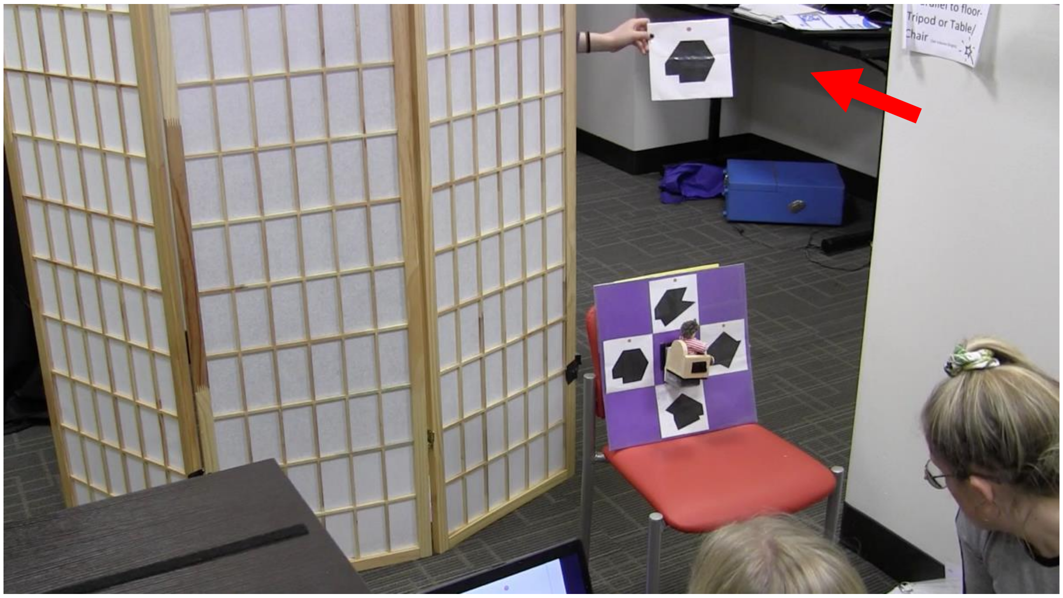
Figure 1: Barrier task set-up



Procedure

At the beginning of the experiment, RA1 framed the barrier task as "playing a game" with RA2. RA1 instructed participants that the objective was to try to get RA2 to pick the shape on the computer screen, as fast as they can. Participants were also informed that RA2 "is facing the shapes the same way that the doll is." Any language that the participant thought was necessary to complete the task was permitted, as long as it was verbally communicated. RA2 would then pick a shape and hold it out from behind the barrier (See Figure 2). RA1 would immediately verify out loud if the shape that was picked by RA2 was correct or incorrect, providing immediate feedback for both the participant and RA2. Each trial began when a shape appeared on the screen, lasted for however long the participant took to speak (there was no time or speech limit on the participant), and ended once RA2 made a shape pick.

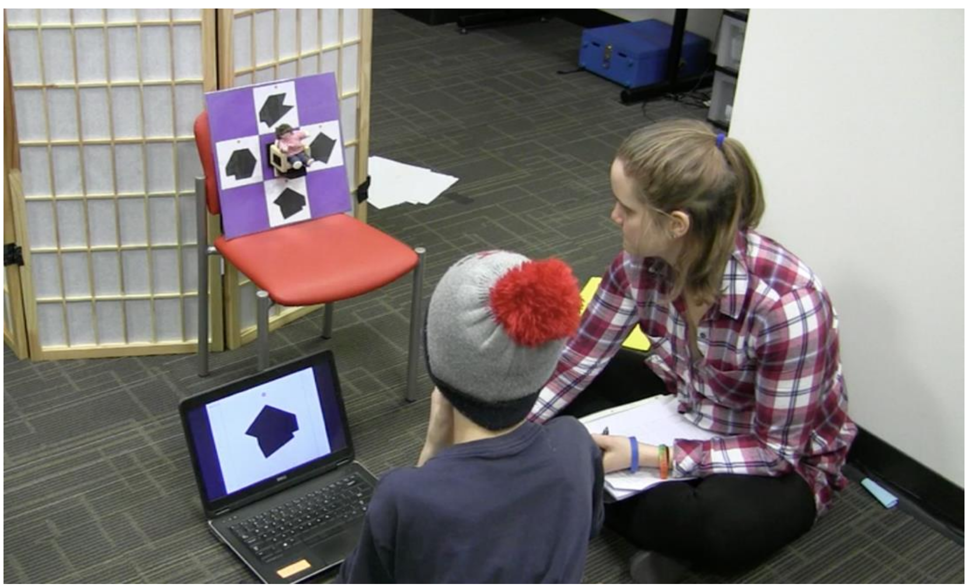
Figure 2: RA2 shape pick



There were two parts to this experiment. In one half on the task, the reference board was placed flat on the floor (See Figure 3). On the other side of the barrier, RA2 was also sitting on the floor surrounded by the four shapes. All the trials with the board on the floor are List A trials (See Appendix A). In the other half of the task, the reference board was placed upright on a chair (See Figure 4). During this part of the experiment, RA2 was sitting in a chair with a one shape under the chair, one pinned to RA2's left, one pinned to RA2's right, and the last one pinned above RA2. All the trials with the board upright are List B trials (See Appendix A).

Figure 3: Reference board on floor (List A trial)

Figure 4: Reference board on chair (List B trial)



Throughout the experiment, the doll was rotated and placed in different configurations. RA2 moved the shapes around to match RA2's perspective of the shapes to the doll's perspective of the shapes. There were a total of eight distinct perspectives, or blocks, with List A and List B having four different perspectives each. For example, in List A, the doll was rotated four times (doll was facing the participant, doll was facing away from the participant, doll was facing in the direction of the participant's right, and doll was facing in the direction of the participant's left) and the shape in front, behind, to the left, and to the right of the doll was different with each rotation. Likewise, the four blocks in List B each had a shape to the right, to the left, above, and below the doll that changed with every doll rotation. Within each block, there were eight trials (i.e., communication task for one shape); two trials per shape. In total, the experiment consisted of 64 trials[3].

Practice trials

At the start of both Lists of trials, participants were given two practice trials, using a practice board and practice shapes (See Figure 5). These practice shapes were behind and in front of the doll in the practice trials of List A and were above and below the doll in practice trials of List B.

Figure 5: Practice trial for List B (upright board)

[3] There were 64 completed trials for all but one included participant, e0124. RA1 allowed the participant to skip five trials with no attempt at communicating with RA2.

Reminders

At the start of a new block, RA1 reminded the participant that RA2 is facing the shapes the same way that the doll is facing the shapes. The same reminder was provided when the participant asked any questions about the task throughout the entire experiment.

RA manipulation

During the first block of the first List (i.e., the first eight shapes of the experiment), if the participant did not use spatial language, RA2 purposely held out an incorrect shape. This was done even if RA2 understood which shape the participant was referencing.

Transcription

Participant videos were uploaded into ELAN, an annotation software, and PARTICIPANT and RA1 tiers were added to transcribe the speech and interactions between the participant and research assistant.

Participant speech was transcribed at the word-level in the PARTICIPANT tier. The participant speech annotations were segmented by pauses lasting longer than one second (whereas pauses less than one second long were indicated by an ellipsis within a single annotation).

RA1 speech consisted of feedback to the participant about whether or not RA2 had selected the correct speech. This speech was not recorded verbatim but instead simply reflected the accuracy of RA2's chosen shape. Thus, RA1 speech was either transcribed as *COR* or *INCOR*.

Transcriptions were all agreed upon by at least two research assistants, as transcriptions were passed from a first transcriber to a second transcriber, and then sometimes to a third. The procedures for transcription were as follows: One research assistant initially transcribed the speech from the video. When they had completed transcribing the entire video, a second research assistant independently reviewed the initial transcriber's transcription while listening to the speech. If there were no discrepancies between what the initial transcriber recorded and what the second transcriber heard, the file was considered final and ready for coding. If the second transcriber noticed moments of discrepancy – i.e., moments where the initial transcriber recorded speech that was different from what the second transcriber heard – they indicated this in the transcription. In these cases, a third transcriber was brought into to review these moments of discrepancy and to make a final decision. In cases where the third transcriber could not decide between the two versions, we established a procedure for consensus transcribing (where all three researchers would listen to the speech together and come to a decision). However, consensus transcribing was never required, as all first-/second-pass discrepancies were decided by the third coder.

Coding

Coding participant speech

In ELAN, all participant speech was coded in two tiers: OBJECTS and TRIAL (See Figure 6). In the OBJECTS tier, codes were assigned to reflect the shapes – numbered *1-4* (Table 2) – that was currently being presented to the participant on the computer screen. The two practice shapes were also coded in the OBJECTS tier, as P1 and P2. Thus, each segment of participant speech was aligned with and identified by the shape the participant was currently describing (See Figure 6). Speech that was unrelated to the shapes was coded as *N/A*. This included irrelevant comments (e.g., "all I want to do is to get this thing over with"), clarification questions directed towards the researchers (e.g., "I'm a little confused like do I tell her? … or"), speech in between sets when there was no shape on the computer screen, and reactions to the shape that the researcher presented from behind the barrier (e.g., "that was close though").

Figure 6: OBJECTS and TRIAL tier in ELAN

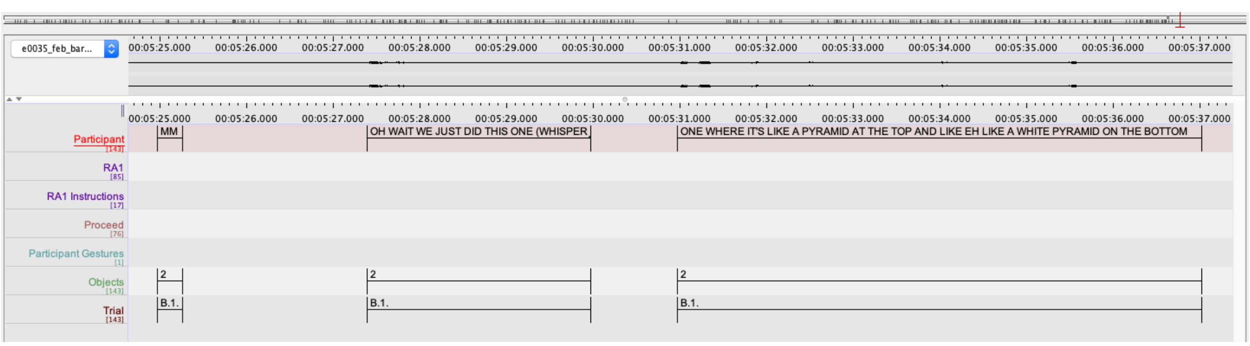
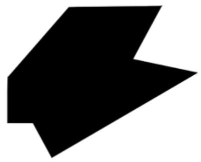
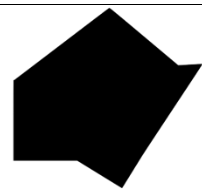
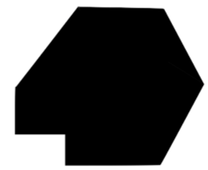


Table 2: OBJECTS code for each shape

1	**2**	**3**	**4**
			

Since each shape appeared 16 times in the context of eight distinct configurations, the TRIAL tier was created to uniquely code speech on each shape. Within this tier, relevant speech (i.e., speech that was used to identify a trial shape) was coded as either *HOR/VER* or with a letter-number combination *A-H.1-2. HOR* and *VER* were used to identify each of the two practice trials when the board was laid flat on the ground (*HOR*) and positioned upright (*VER*). For experimental trials, the letters *A-H* were used to indicate the eight different blocks, which correspond to eight different positions of each shape relative to the doll. Codes beginning with *A-D* indicates horizontal-board trials, and codes *E-H* indicates vertical-board trials. For example, codes beginning with *A* reflect trials in the A block, during which the board is flat on the ground and the doll is facing the participant. As explained in procedures, each shape appears twice within each block, so the number *1-2* was added to the TRIAL code to indicate whether the child is describing the shape for the first or second time within the block. The OBJECTS tier code appended to the TRIAL tier code produces a unique identifier for each of the 64 trials (Table 2). For example, *A.1.1* identifies the first time shape 1 is presented to participants in the block where the board is flat on the ground and the doll is facing the participant. All segments coded as *N/A* in the OBJECTS tier was also coded as *N/A* in the TRIAL tier. (See Appendix B for all trial codes).

Coding for shape description communication strategies

A coding scheme was created for analysis of all the communication strategies employed by every participant included in the overall study. Of this larger scheme, a subset of these strategy codes were used for this study. These included only the trials when participants described the shapes to RA2 in some way, rather than identifying the shape by referencing its position to RA2 or the doll. Through observation of the data, two subcategories emerged under the Object related strategies: *Name* and *Describe.* Trials were coded with *Name* if the participant communicated about the shape with a label. For example, “it’s the mushroom” or, simply, “mushroom”. Under the *Describe* communication strategy, three further subcategories were formed: *Geometry*, *Resemblance*, and *Other*.

Geometry

Trials were coded with the *Geometry* strategy if the participant described the shapes in reference to the shape’s lines and/or angles. Special inclusions under the *Geometry* strategy included descriptions where the participant provided specific instruction to RA2 on how to draw the outline of the shape (e.g., “…whenever that edge goes down, it’s like a pointy edge another one goes down the ramp from the dot and it’s like a side that’s like flat and then it goes in ninety degrees and then comes out”), used geometric shape names (e.g., “it has a triangle at the top”, “it’s shaped like a hexagon”), or identified the shape as being shaped in the same way as an alphabetic letter (e.g., “it’s V-shaped”, “it’s L-shaped”).

Resemblance

Trials coded with the *Resemblance* strategy involved descriptions about the shape looking similar to a real-world object (e.g., "it looks like a mushroom"). These *Resemblance* trials were further specified as either whole object descriptions (e.g., “it looks like a hexagon with a piece cut out”) or partial object descriptions (e.g., "the top part looks like a spike"). While the use of geometric shape names was coded under *Geometry*, instances where the participant stated that the shape “looked like” a geometric figure were considered *Resemblance*.

Other

Trials coded with the *Other* strategy comprised of nondiscriminatory features of the shape in the description. These were the details that were broadly shared among more than one of the shapes (e.g., mention of the shape being black, having eight sides, or having an orange dot above the shape).

In ELAN only one strategy code annotation was made for each trial. Therefore, *Combined* was an additional subcategory that applied to trials where the participant used more than one description strategy within a single trial.

Analysis

Overall accuracy and strategy effectiveness

Accuracy

Participant accuracy trial-to-trial was converted to binary (0 vs. 1) scores: Participant that had been coded as *COR* in ELAN were scored as 1 and *INCOR* as 0. Then, we summed participant scores and calculated a proportion of correct trials out of the total number of trials[4] in Sets 2 through 8. We excluded Set 1 from accuracy calculations because – as explained earlier in the methods – during Set 1, RA1 intentionally misinterpreted child descriptions if the child did not use spatial language. This was done to encourage children to switch their strategy from object description to spatial description (as the original goal of the study was to compare the accuracy of spatial language between groups). Thus, accuracy in Set 1 is misleading, as the RA was not always a cooperative communication partner. After Set 1, the RA made every effort to interpret child descriptions as the child intended. Proportions were then compared between groups using a two-tailed, heteroscedastic *t*-test.

Effective strategy switching and maintenance

We operationalized 'effective strategy use' as participant's tendency to switch their strategy after an *INCOR* trial and, vice versa, to maintain their strategy after a *COR* trial. In this way, effective strategy use reflected the participant's recognition that their previous strategy was

[4] We used proportions – rather than sums – to compare accuracy between groups and across participants, because not all participants had the same number of trials. This was due to rare instances where RA1 would allow the participant to skip a shape without an attempt at doing the task.

effective (and therefore should be maintained) or was not effective (and therefore should be replaced by a new method).

To quantify effective strategy use, a 1 was assigned when the general type of strategy changed in the following trial after an *INCOR* trial (i.e., the participant switched type of communication strategy after negative feedback was given) and when the strategy codes remained the same or referentially shortened (e.g., the participant says, “this one looks like an arrow”, and in the very next trial of the same shape, the participant says, “the arrow”) *COR* trial (i.e., the participant did not switch strategies after positive feedback was given). An ‘effective strategy quotient’ was then calculated by summing the 1s for each child and dividing this total by the total number of inter-trial turns. We compared effective strategy quotients between groups also by a two-tailed, heteroscedastic *t*-test.

Object-description trials only

Referential shortening

“Referential shortening demonstrates implicit adherence to discourse rules related to common ground” (De Marchena and Eigsti, 2016). A similar method to the study by De Marchena and Eigsti was used. However, the number of words used for any given trial varied widely due to individual differences. Therefore, the number of words spoken was controlled by converting the data to standard deviations from the participant’s mean. This made the participant’s referential shortening relative to each participant and allowed for same-group and between-group comparisons.

Common ground establishment and maintenance

To determine whether the participants established and then maintained common ground when describing/labeling shapes, a similar calculation was used to the ‘effective strategy quotient’ described above in *Analysis* section 1.2 for shape-description trials. For this analysis, only shape-description trials were considered, and the participant’s description of each shape was compared during repeated presentations of that shape. For example, the way the child described Shape 1 was compared to the first time she saw it to the second time she saw it, etc.

Similar to the way effective strategy use (*Analysis* section 1.2) was determined, RA performance/feedback on the previous trial was used, combined with comparing the child's strategy on the current trial to that of the previous trial, to gain a measure of how motivated the child was to establish common ground and then whether the child maintained common ground once (if) it was established. These values resulted in four 'common-ground quotients' (one for each shape) per participant.

Therefore, to capture the participant's motivation to *establish* common ground, the participant received a 1 if they changed their description of a shape after an *INCOR* trial for the same shape. For example, a child describes Shape 1 as "looking like a horse head" on one trial. If the RA picked the incorrect shape, and then the child described it as "a mushroom" on the next trial, the child would receive a 1. If the child again called the shape "the horse head" on the subsequent trial, they would earn a 0.

Similarly, to capture the participant's ability to maintain common ground, they would receive 1 when the participant's description or label for a shape remained consistent after a *COR* trial for the same shape. If the participant changed their description after a *COR* object-description trial, they earned a 0. A description was considered the same whether the child described it (e.g., "it looks like a hexagon with a bite taken out of it") or applied a label (e.g., "it's the hexagon"), as long as crucial aspects of the description remained present (e.g., "hexagon").

We then summed the 1s for each shape and divided this number by the number of trials for that shape where the child used shape description/labeling. Thus, each participant was assigned four common-ground quotients (one for each shape).

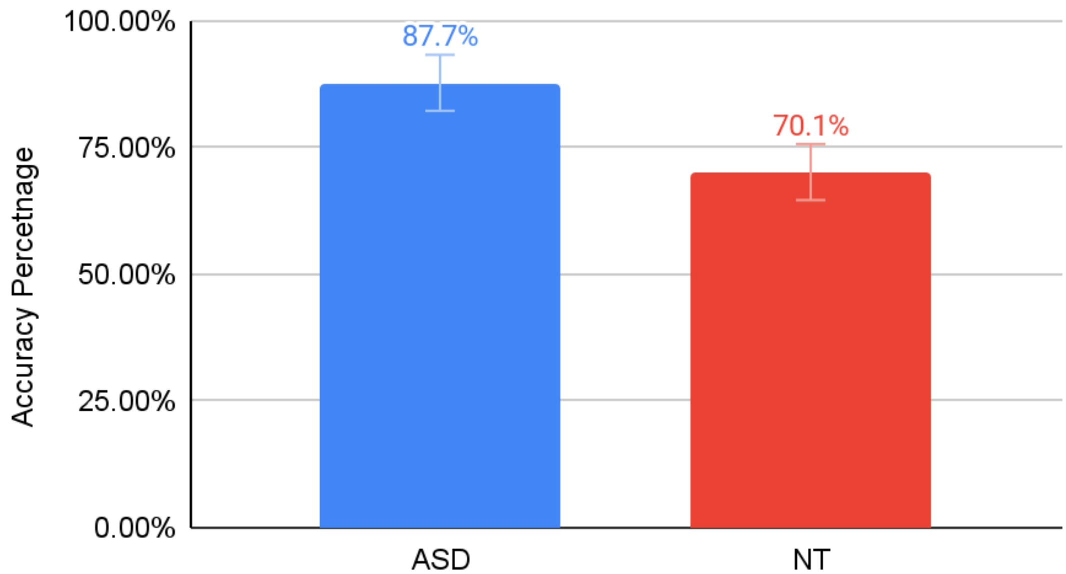
To compare common ground establishment/maintenance between groups, a repeated-measures 2x4 ANOVA was used to determine whether group membership and shape significantly predicted common-ground quotients. Shape was included as a predictor as it is possible that certain shapes (e.g., Shape 4 was frequently and quickly identified as a "hexagon" by many participants) allowed for easier common-ground establishment than others.

RESULTS

Accuracy

Out of the 64 experimental trials (not including practice trials), participants in the ASD group had RA2 pick the correct shape, on average, 87.7% of the time, while the NT group had an accuracy average of 70.1% (Figure 7). A two-tailed, heteroscedastic t-test revealed no significant difference between the two groups (t = 1.7048, df = 9.5718, p = 0.1204).

Figure 7: Accuracy percentage by group



Effective Strategies

After every trial except for the last, participants had an opportunity to switch or maintain a communication strategy. Figure 8 shows the average number of each efficacy type (effective switches, effective maintenance, ineffective switches, and ineffective maintenance) done by each group. We used a repeated-measures 2x4 (Group x Efficacy Type) ANOVA to compare the use of effective vs. ineffective maintenance vs. switching between groups. It revealed no significant effect of group ($F_{(1,13)} < 0$, $p = 0.999$, $\eta^2_G = 6.64e\text{-}33$) but significant effect of efficacy type ($F_{(3,39)} = 12.71$, $p < 0.001$, $\eta^2_G = 4.94e\text{-}01$). However, there was no interaction between group and efficacy type ($F_{(3,39)} = 1.72$, $p = 0.179$, $\eta^2_G = 0.12$).

Out of the 63 opportunities to switch or maintain, the ASD group, on average, effectively switched and maintained their type of strategy 69.2% of the time. The NT group had a slightly lower average of 61.5% (Figure 9). Groups were not significantly different according to a two-tailed, heteroscedastic *t*-test comparing effective strategy proportions ($t = 0.81393$, $df = 11.609$, $p = 0.4321$).

Considering only the effective switches and maintenance, Figure 10 displays the proportion of switches and the proportion of maintenance within these trials. A two-tailed, heteroscedastic t-test found that the proportion of effective strategy maintenance and switches were not significantly different between groups ($t = 1.7679$, $df = 8.7221$, $p = 0.1119$).

Figure 8: Average number of effective and ineffective switching/maintenance by group

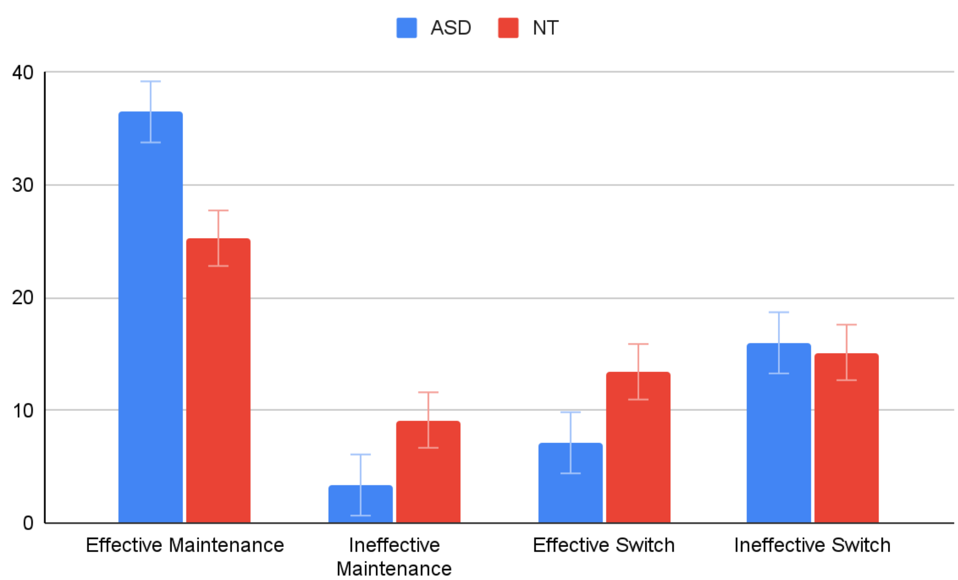


Figure 9: Effective strategy percentage by group

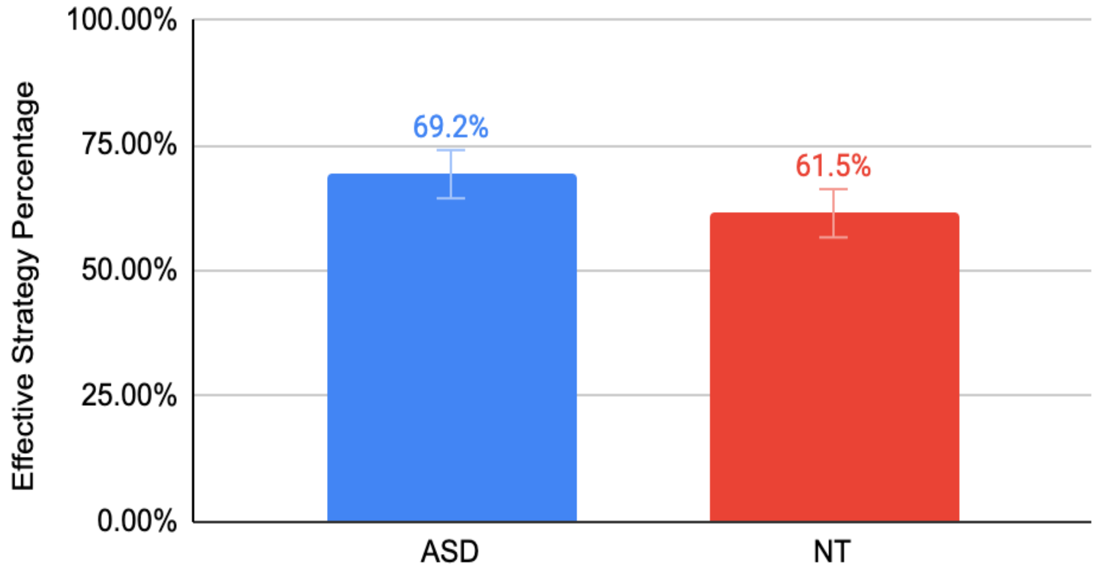
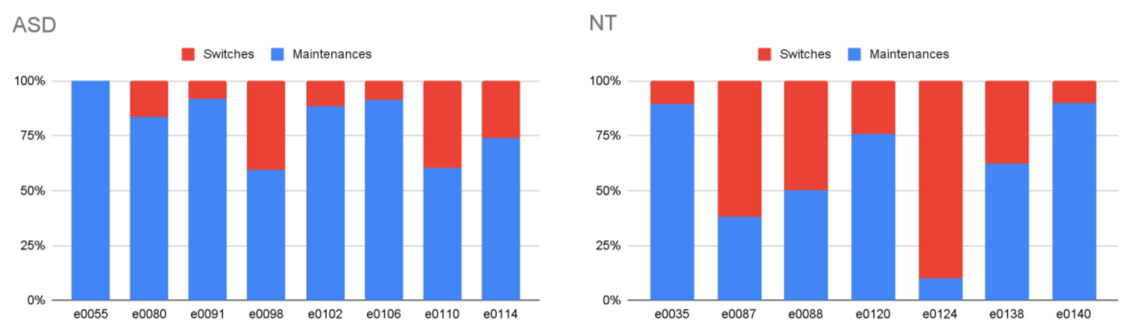


Figure 10: Proportions of effective maintenances/switches. Participants separated by group.



Same-Shape Description Trials

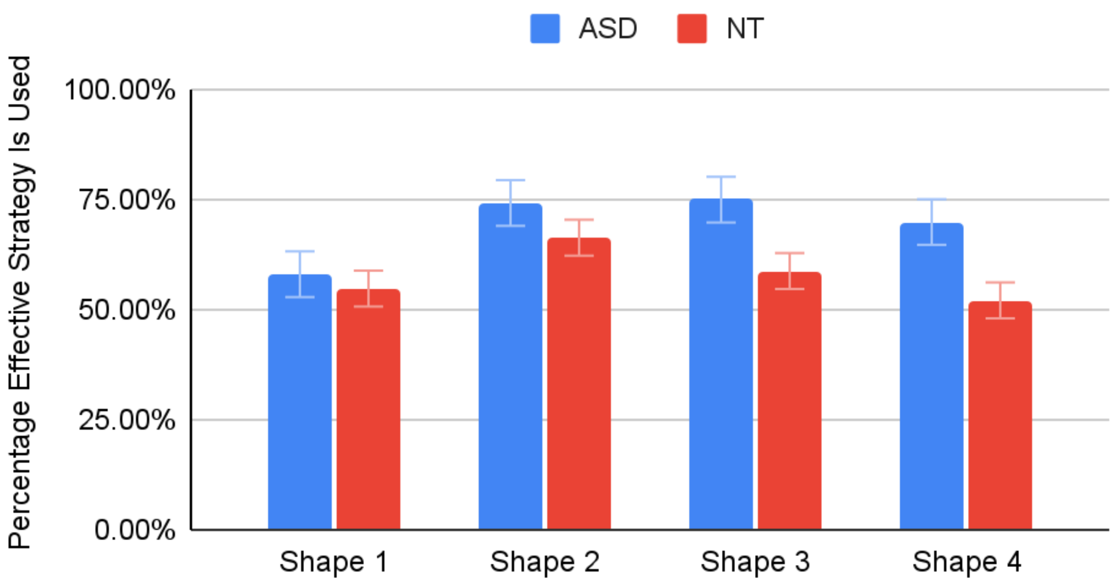
Of the 15 participants, a subgroup of ten had a sufficient number of object-description trials (i.e., subgroup participants used a type of description strategy at least two times for all four shapes) – 4 ASD and 6 NT. Table 3 and Figure 11 summarize the results of effective switches and maintenance of each shape by group. Repeated-measures 2x4 (Group x Object) ANOVA to predict effective strategy quotients. This test revealed no significant effect of group ($F_{(1,8)} = 1.67$, $p = 0.232$,

$\eta^2_G = 0.08$), no significant effect of object ($F_{(3,24)} = 1.04$, $p = 0.393$, $\eta^2_G = 0.07$), and no significant interaction between these two factors ($F_{(3,24)} = 0.33$, $p = 0.805$, $\eta^2_G = 0.02$).

Table 3: Average effective strategy proportions for each shape in each group. Numbers in parentheses represent standard error.

	Shape 1	***Shape 2***	***Shape 3***	***Shape 4***
ASD	58.1% (5.22%)	74.2% (16.24%)	75.0% (11.85%)	69.9% (6.74%)
NT	54.8% (5.47%)	66.3% (9.70%)	58.8% (11.05%)	52.1% (6.07%)

Figure 11: Effective strategy proportion by group on same-shape object trials



Referential Shortening

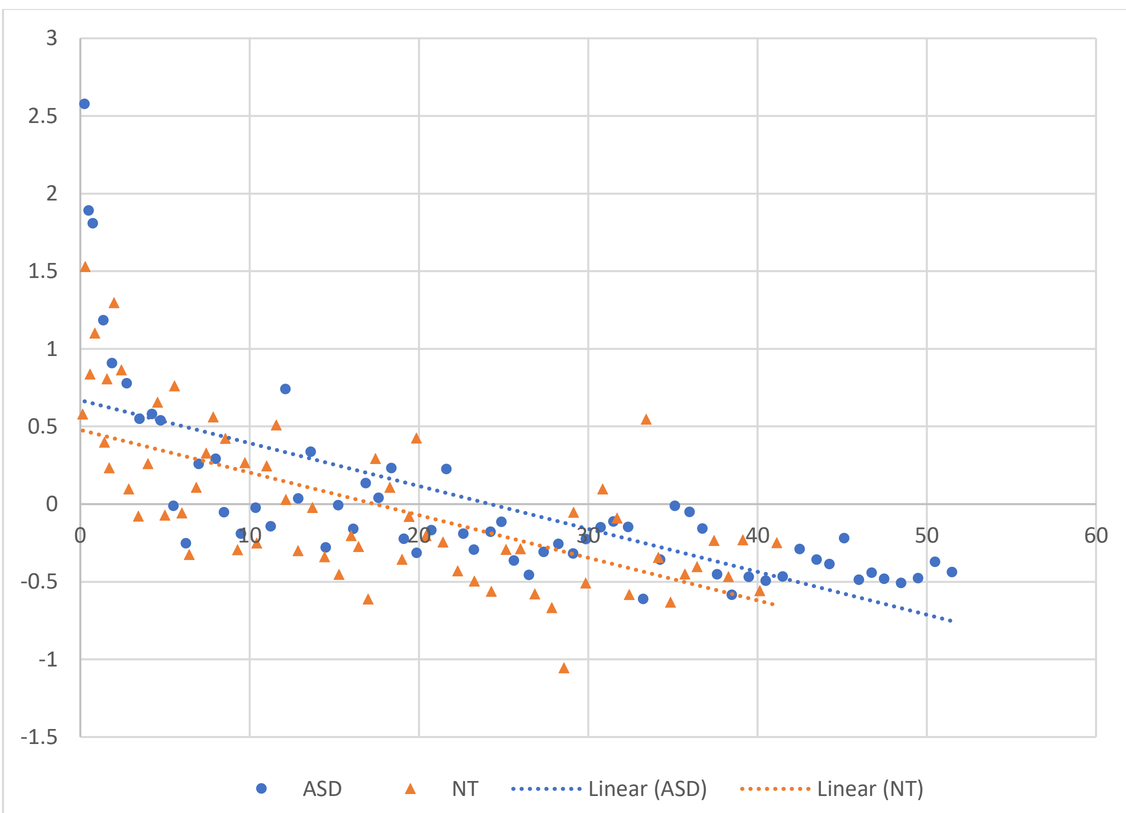
We used a linear mixed-effects model to determine whether trial (from 1 to 64) and diagnosis (ASD vs. NT) affected the standardized word count.

We find a significant main effect of *Trial*, with fewer words (relative to each participant) more likely to occur in later trials as compared to earlier ones (β = -0.025, SE = 0.002, t(943) = -11.59, $p < 0.0001$).

There was also a marginal interaction between Trial and Group, where NT individuals showed a more positive (i.e., less steep negative slope) in the number of words used across trials (relative to themselves) as compared to participants with ASD (β = 0.006, SE = 0.003, t(943) = 1.880, *p* = 0.061). See Figure 12.

The main effect of Group is not significant for this model, but it is also not easily interpretable, as the dependent variable is standardized. To account for this, we ran a second model with actual word count (not standardized) as the dependent variable, and there is still not a significant main effect of Diagnosis (β = 8.01, SE = 5.66, t(14.81) = 1.42, *p* = 0.177), showing that there was not a significant difference in words used by each group, overall, although the NT group does use more words on average (ASD m = 11.3(14.3); NT m = 20.0(19.3)).

Figure 12: ASD and NT group sum accuracy by word count (standard deviation from mean).



DISCUSSION

The current study initially hypothesized that participants with ASD would perform worse than their NT peers at effectively and efficiently adapting to the needs of their communication partner. However, our results present a picture where participants with ASD were as effective at communicating in this study as their NT peers. We discuss each of our findings in the following.

Accuracy

Our first set of findings indicate equal task accuracy between both groups. Many previous studies have found that individuals with ASD perform poorly in referential communication tasks (Malkin et al., 2018; Schuh et al., 2016; Dahlgren and Dahlgren Sandberg, 2008; Volden et al., 1997), but our findings indicate that both the children with and without ASD were similarly able to communicate in a way that resulted in their communication partner selecting the target shape/object. In fact, the ASD group achieved higher accuracy scores, on average, than their NT peers did. Although these differences weren't significant, high accuracy in the ASD group suggests that these participants were effective at using language that allowed their partner to identify the targeted shape.

It is possible that the surprising (surprising, in light of previous findings) success of the ASD group on this task is attributable to the feedback system we used. Our experiment employed a simplified version of communication where the provided feedback was a straightforward binary response (i.e., "yes" or "no"). Thus, participants received relatively explicit signals about the success of their communication attempt, which has not been implemented in previous research. For example, Nadig et al. (2015) found that individuals with ASD were significantly less successful at communicating reference in a barrier task, similar to ours. However, participants in the Nadig et al. (2015) study were only able to determine their lack of success from questions or statements with upward inflection (success was directly indicated by the research assistants, who were instructed to state, "okay" or "got it"). Therefore, it is possible that individuals with ASD can communicate reference effectively, and can change their strategy to meet listener needs, if their listener is explicit about their ability to understand the message.

Further, not only is explicit “yes/no” feedback more direct than what has been used in previous referential communication studies, but it is incredibly more obvious than the type of signals that an interlocutor would be more likely to use in a natural interaction (e.g., facial expressions signaling confusion, clarification questions, topic changes). It is known that individuals with ASD often struggle to interpret facial expressions and other non/extralinguistic signals, like prosody (Loth et al., 2018; Uljarevic and Hamilton, 2013; Harms et al., 2010; Clark et al., 2008; Lopez et al., 2004). By removing participants' need to interpret such signals (and even removing the ability to observe facial expressions, since their interlocutor was behind a barrier), we may have made groups better matched in recognizing their partners' needs and mental states[5]. Had our task involved more complex parameters, such as providing open-ended feedback (vs. our “yes”/“no” feedback system), these previously stated differences in referential communication could have been reflected in our results as well.

It is possible that our study could be providing a more precise depiction of referential communication in children with ASD. Previous studies that investigated the referential ability of children with ASD with similar tasks as ours challenged the participants’ consideration of the listener’s visual perspective and selection of an appropriate referring expression (Nadig et al., 2009; Fukumura, 2016). Both studies found a diminished performance in the ASD group. In the current study, the participants selected for this investigation only used descriptive strategies (thereby bypassing any need to consider the partner’s visual perspective). Therefore, since our study does not find differences between groups, the contrast in performance may have only been with consideration of another’s visual perspective. Our results confirm that children with ASD can select appropriate referring expressions, supporting the claim that children with ASD do not struggle with

[5] This was likely helped by the fact that the participant groups are indeed well matched -- earning equivalent scores on language and cognitive tests.

referential communication but instead with the extraneous aspects of communication, such as executive functioning, expression recognition (Loth et al., 2018; Uljarevic and Hamilton, 2013; Harms et al., 2010; Clark et al., 2008; Lopez et al., 2004), and central coherence (Happé & Frith, 2006).

Effective Strategy Switching and Maintenance

Our second set of findings shows evidence that both ASD and NT children are effective at maintaining or switching their communication strategy based on the success or failure of a previous attempt to communicate reference. One possibility for this is that the ASD participants did in fact use an ineffective reference to describe the objects/shapes but ultimately achieved a similar effective strategy quotient as the NT group simply due to their consistency throughout the entirety of the task (i.e., maintenance of their own strategies) and the flexibility of the communication partner.

While participants in the ASD group earned equivalent effective strategy quotients to NT peers, we actually do not have evidence of the ASD children doing much *adapting* to their listener (again, they earned high scores by simply continuing a previous strategy). Not only do we suspect that strategy changes would be more challenging for them due to general tendencies to prefer sameness, but strategy changes also depend on executive functioning (EF) more than maintenance does. The cognitive demands for *strategy switching* likely involve several EF components, including: (1) attention shifting, since the child must shift attention to the listener's signals; (2) working memory, since the child must keep track of unsuccessful previous strategies in order to know what to avoid going forward; and (3) inhibitive skills, since the child must suppress the old strategy to resourcefully think of a new strategy that may improve the listener's understanding. Research has found evidence that individuals with ASD show EF differences, not only compared to NT peers but also to peers with disorders typically associated with EF challenges, like ADHD (Corbett et al., 2009; Nilsen and Graham, 2009; Weismer et al., 2018). Corbett et al. (2009) evaluated the EF profiles of children with ASD, children with ADHD, and NT children, revealing greater impairment in inhibition, working memory, cognitive-flexibility/shifting, and vigilance (attention) in the ASD group as compared to both other groups. Weismer et al. (2018) confirmed these results in a study examining the association between EF skills and language abilities in children with ASD and NT children. Their study found similar results of significant differences in all components of EF for the ASD group as well as an association between these EF weaknesses and language weaknesses. In fact, the inhibition of one's own thoughts/perspectives has been identified as a factor for referential communication weaknesses in this population (Weismer et al., 2018) since inhibitive skills are arguably requisite in being able to understand the thoughts or perspectives of others.

We also examined the *type* of helpful adaptations (i.e., maintenance vs. switching) between groups. The majority of ASD participants' high effective strategy quotients were due to a large number of trials where the participant appropriately *maintained* their previous strategy. In contrast, the NT group had a similar task accuracy but with a greater number of trials on average that implemented an effective strategy switch. We interpret this finding as reflecting the fact that participants in the ASD group identified an effective strategy early on (or a mutually understood way with their partner). This meant that these participants could achieve subsequent success simply by preserving the strategy they had been using. Thus, participants with ASD were not required to change their strategy in order to achieve communication success and in order to achieve a high effective strategy quotient. We are unable to determine whether participants in the ASD group would know to *change* strategies to adapt to listener needs, but we do suspect that this skill would be harder for them than strategy maintenance. This is supported by the DSM-5 criteria for ASD, which states that individuals with ASD have highly restricted, fixated interests and demonstrate cognitive inflexibility. Perseverative communicative patterns in individuals with ASD further supports our results of the greater proportion of maintenance trials in the ASD group, indicating even more so that the high effective strategy quotient may be due to the *listener* adapting to the ASD participant.

Referential Shortening

Our third set of findings reveals that both groups referentially shorten their responses over time at similar rates. Referential shortening measured by word count suggests communicative efficiency, so both groups demonstrated an awareness and an ability to become more efficient with their communication over time. As described in the above sections, participants with ASD were quite good at determining an effective strategy early on in the task. This meant that they could afford to use fewer words to communicate as effectively as the individuals in the NT group, who switched strategies more often which involved introducing new (and, therefore, lengthy) explanations to their listener.

These findings are contrary to previous research which reports that children with ASD are less efficient with their responses than NT children, providing more redundant and irrelevant information (Volden et al., 1997; Dahlgren and Dahlgren Sandberg, 2008). Further, De Marchena and Eigsti (2016) specifically compared referential shortening effects between ASD and NT groups and found that the ASD group did not show evidence of referential shortening. However, all the previous research has failed to measure referential shortening of responses in dynamic discourse (i.e., where the speaker's message is influenced by the listener's response or understanding). De Marchena and Eigsti's study also only compared referential shortening effects between conditions: one situation when they knew their listener did not share common ground with them and another when they did. The methods we used in our study more closely reflects actual communication, where speakers and listeners are *both* interacting and cooperatively creating effective and efficient responses over time. While the responses from the ASD participants of our study were not free from redundancies and irrelevancies, the decreasing word count over time suggests that children with ASD are aware of at least some pragmatic tendencies for efficient communication.

Study Limitations and Future Directions

The current study has a few limitations that present the need for further investigation. For one, our participants were selectively sampled, and – by doing so – we may have selected participants whose communication patterns are not representative of their group, overall. We determined shape descriptions to be an inefficient way of performing this task, since trials with shape descriptions tended to have high word counts and often resulted in inaccurate picks. Therefore, we chose participants who used this as a strategy in order to maximize our observation for possible improvements in communicative efficiency over time. While this allowed for a more robust analysis, it limits our findings to only a few types of individuals out of a larger set of NT and ASD individuals. Specifically, by selecting participants who initially used less effective communication strategies, we may have ended up with participants (in both groups) who were less skilled in communication effectiveness, overall. This may help explain why participants in the NT group performed surprisingly poorly, in terms of their strategy switching, accuracy, and speed of referential shortening. Another limitation of using this sample is that it left us with a smaller sample size. Finally, as always when testing participants with autism who test as having at least typical language and intelligence, results are (likely) not generalizable to the ASD population, at large.

Clinical Implications

While our results did not follow our initial predictions, the ASD group's success in our study provides us with potential implications for therapeutic methods. Based off our findings, treatment targeting referencing in discourse for individuals with ASD should use direct, immediate, and straight-forward feedback. Our results indicate that when these children with ASD are provided this type of feedback, they are successful in their interactions, both referentially and pragmatically. Relatedly, clinicians may consider incorporating self-advocation goals into treatment plans for individuals with ASD. Since we find that individuals with ASD can appropriately reference and participate in social discourse in the context of receiving clear and contemporaneous feedback, advocating for oneself by asking for this type of explicit response is a reasonable accommodation that these individuals can realistically request from others.

Conclusions

In summary, we analyzed the effectiveness and efficiency of communication strategies used by ASD and NT children through a method that monitored maintenance and switching (or adaptations) in strategy according to task accuracy. We found that both groups were equally effective at adapting their communication throughout the task. We also studied the pragmatic aspect of referential shortening and found it to be present in both group and similar rates. Our study presents a deeper understanding of referential communication abilities in children with ASD due to our task methodology and measures.

Appendix A

Order of Shape Presentation in List A and List B

List A	**List B**
A.1.1	E.1.1
A.1.4	E.1.4
A.1.2	E.1.2
A.1.3	E.1.3
A.2.4	E.2.4
A.2.1	E.2.1
A.2.3	E.2.3
A.2.2	E.2.2
B.1.2	E.1.2
B.1.3	F.1.3
B.1.4	F.1.4
B.1.1	F.1.1
B.2.3	F.2.3
B.2.2	F.2.2
B.2.1	F.2.1
B.2.4	F.2.4
C.1.4	G.1.4
C.1.1	G.1.1
C.1.3	G.1.3
C.1.2	G.1.2
C.2.1	G.2.1
C.2.4	G.2.4
C.2.2	G.2.2
C.2.3	G.2.3
D.1.3	H.1.3
D.1.2	H.1.2
D.1.1	H.1.1
D.1.4	H.1.4
D.2.2	H.2.2
D.2.3	H.2.3
D.2.4	H.2.4
D.2.1	H.2.1

Appendix B

Appended OBJECT and TRIAL Tier Codes

	P1	*P2*	*1*	*2*	*3*	*4*
HOR.	HOR.P1	HOR.P2				
VER.	VER.P1	VER.P2				
A.1.			A.1.1	A.1.2	A.1.3	A.1.4
A.2.			A.2.1	A.2.2	A.2.3	A.2.4
B.1.			B.1.1	B.1.2	B.1.3	B.1.4
B.2.			B.2.1	B.2.2	B.2.3	B.2.4
C.1.			C.1.1	C.1.2	C.1.3	C.1.4
C.2.			C.2.1	C.2.2	C.2.3	C.2.4
D.1.			D.1.1	D.1.2	D.1.3	D.1.4
D.2.			D.2.1	D.2.2	D.2.3	D.2.4
E.1.			E.1.1	E.1.2	E.1.3	E.1.4
E.2.			E.2.1	E.2.2	E.2.3	E.2.4
F.1.			F.1.1	F.12	F.13	F.14
F.2.			F.2.1	F.2.2	F.2.3	F.2.4
G.1.			G.1.1	G.12	G.1.3	G.1.4
G.2.			G.2.1	G.2.2	G.2.3	G.2.4
H.1.			H.1.1	H.1.2	H.1.3	H.1.4
H.2.			H.2.1	H.2.2	H.2.3	H.2.4

References

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

Baron-Cohen, S. (2000). Theory of mind and autism: A fifteen year review. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 3-20). Oxford University Press.

Barr, D. J., & Keysar, B. (2006). Perspective Taking and the Coordination of Meaning in Language Use. In *Handbook of Psycholinguistics* (2nd ed., pp. 901-938). Elsevier Science & Technology. 10.1016/B978-012369374-7/50024-9

Clark, H. H., Murphy, G. L., & Kintsch, W. (1982). Audience design in meaning and reference. In J.-F. Le Ny (Ed.), *Language and comprehension* (pp. 287-299).

Clark, T. F., Winkielman, P., & McIntosh, D. N. (2008). Autism and the extraction of emotion from briefly presented facial expressions: Stumbling at the first step of empathy. *Emotion*, *8*(6), 803–809. 10.1037/a0014124

Corbett, B. A., Constantine, L. J., Hendren, R., Rocke, D., & Ozonoff, S. (2009). Examining executive functioning in children with autism spectrum disorder, attention deficit hyperactivity disorder and typical development. *Psychiatry Research*, *166*(2-3), 210-222. PMC. 10.1016/j.psychres.2008.02.005

Dahlgren, S., & Dahlgren Sandberg, A. (2008). Referential communication in children with autism spectrum disorder. *Autism*, *12*(4), 335-348. 10.1177/1362361308091648

De Marchena, A., & Eigsti, I.-M. (2016). The art of common ground: emergence of a complex pragmatic language skill in adolescents with autism spectrum disorders. *Journal of Child Language*, *43*(1), 43-80. 10.1017/S0305000915000070

Happé, F & Frith, U. (2006). The Weak Coherence Account: Detail-focused Cognitive Style in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5-25. 10.1007/s10803-005-0039-0

Fukumura, K. (2016). Development of Audience Design in Children With and Without ASD. *Developmental Psychology*, *52*(1), 71–87. APA PsycArticles. https://doi.org/10.1037/dev0000064

Harms, M. B., Martin, A. & Wallace, G. L. (2010). Facial Emotion Recognition in Autism Spectrum Disorders: A Review of Behavioral and Neuroimaging Studies. *Neuropsychol Rev 20*, 290–322. 10.1007/s11065-010-9138-6

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*(1), 91-117. 10.1016/0010-0277(96)81418-1

John, A. E., Rowe, M. L., & Mervis, C. B. (2009). Referential Communication Skills of Children with Williams Syndrome: Understanding When Messages are Not Adequate. *American journal on intellectual and developmental disabilities*, *114*(2), 85-99. 10.1352/2009.114.85-99

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman brief intelligence test: KBIT-2*. Bloomington, MNPearson.

Krauss, R. M., & Weinheimer, S. (1964). Changes in reference phrases as a function of frequency of usage in social interaction: A preliminary study. *Psychonomic Science*, *1*(5), 113-114. APA PsycInfo. https://doi.org/10.3758/BF03342817

Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology*, 4(3), 343-346.

Landa, R. (2000). Social language use in Asperger syndrome and high-functioning autism. In A. Klin, F. R. Volkmar, & S. S. Sparrow (Eds.), *Asperger syndrome* (pp. 125-155). The Guilford Press.

Loth, E., Garrido, L., Ahmad, J., Watson, E., Duff, A., & Duchaine, B. (2018). Facial expression recognition as a candidate marker for autism spectrum disorder: how frequent and severe are deficits?. *Molecular Autism*, *9*(7). 10.1186/s13229-018-0187-7

Lopez, B., Donnelly, N., Hadwin, J., & Leekam, S. (2004). Face processing in high-functioning adolescents with autism: Evidence for weak central coherence. *Visual Cognition, 11*(6). 673-688. 10.1080/13506280344000437.

Lord, C., DiLavore, P. C., Gotham, K., Guthrie, W., Luyster, J., Risi, S., & Rutter, M. (2012). *Autism Diagnostic Observation Schedule: ADOS-2*. Los Angeles, CA: Western Psychological Services.

Malkin, L., Abbott-Smith, K., & Williams, D. (2018). Is verbal reference impaired in autism spectrum disorder? A systematic review. *Autism & Developmental Language Impairments*, *3*, 1-24. https://doi.org/10.1177/2396941518763166

Marvin, R. S., Greenberg, M. T., & Mossler, D. G. (1976). The Early Development of Conceptual Perspective Taking: Distinguishing among Multiple Perspectives. *Child Development*, *47*(2), 511-514. 10.2307/1128810

Nadig, A., Shivani, S., & Sasson, M. (2015). Global Similarities and Multifaceted Differences in the Production of Partner-Specific Referential Pacts by Adults with Autism Spectrum Disorders. *Frontiers in Psychology*, *6*.

Nadig, A., Vivanti, G., & Ozonoff, S. (2009). Adaptation of object descriptions to a partner under increasing communicative demands: a comparison of children with and without autism. *Autism Research*, *2*(6), 334-347. https://doi.org/10.1002/aur.102

Nilsen, E. S., & Graham, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology*, *58*(2), 220-249. https://doi.org/10.1016/j.cogpsych.2008.07.002

Rutter, M., Bailey, A., & Lord, C. (2003). *Social Communication Questionnaire (SCQ)*. Los Angeles, CA: Western Psychological Services.

Schuh, J. M., Eigsti, I.-M., & Mirman, D. (2016). Discourse comprehension in autism spectrum disorder: Effects of working memory load and common ground. *Autism Research*, *9*(12), 1340-1352. Epub. 10.1002/aur.1632

Tager-Flusberg, H., Paul, R., Lord, C., & Donald, C. J. (2005). Language and Communication in Autism. In F. R. Volkmar, R. Paul, & A. Klin (Eds.), *Handbook of autism and pervasive developmental disorders: Diagnosis, development, neurobiology, and behavior* (pp. 335-364). John Wiley & Sons, Inc.

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: a formal meta-analysis. *Journal of autism and developmental disorders*, *43*(7), 1517–1526. 10.1007/s10803-012-1695-5

Volden, J., Mulcahy, R. F., & Holdgrafer, G. (1997). Pragmatic language disorder and perspective taking in autistic speakers. *Applied Psycholinguistics*, *18*(2), 181-198. Cambridge University Press. https://doi.org/10.1017/S0142716400009966

Weismer, S. E., Kaushanskaya, M., Larson, C., Mathée, J., & Bolt, D. (2018). Executive Function Skills in School-Age Children With Autism Spectrum Disorder: Association With Language Abilities. *Journal of Speech, Language, and Hearing Research*, *61*(11), 2641–2658. 10.1044/2018_JSLHR-L-RSAUT-18-0026

Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals–Fifth Edition (CELF-5)*. Bloomington, MN: NCS Pearson.

Yule, G. (1996). Reference and inference (H. G. Widdowson, Ed.). In *Pragmatics*. Oxford University Press.