A NOVEL GAZE BASED MEASURE OF RECEPTIVE VOCABULARY AND CATEGORY KNOWLEDGE IN TODDLERS AT A HEIGHTENED FAMILIAL RISK FOR AUTISM SPECTRUM DISORDER

by

Kathryn M. Hauschild

Bachelor of Science, University of Massachusetts, 2011

Master of Science, University of Pittsburgh, 2014

Submitted to the Graduate Faculty of

The Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

THE KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Kathryn M. Hauschild

It was defended on

May 23, 2018

and approved by

Mark S. Strauss, PhD, Associate Professor, Department of Psychology

Jana M. Iverson, PhD, Professor, Department of Psychology

Melissa Libertus, PhD, Assistant Professor, Department of Psychology

Diane L. Williams, PhD, CCC-SLP, Professor, Department of Communication Sciences & Disorders, The Pennsylvania State University

Dissertation Chair: Mark S. Strauss, PhD, Associate Professor, Department of Psychology

A NOVEL GAZE BASED MEASURE OF RECEPTIVE VOCABULARY AND CATEGORY KNOWLEDGE IN TODDLERS AT A HEIGHTENED FAMILIAL RISK FOR AUTISM SPECTRUM DISORDER

Kathryn M. Hauschild, PhD

University of Pittsburgh, 2018

Toddlers with autism spectrum disorder (ASD) may have language, social, and/or communicative deficits (i.e. gaze aversion and reduced understanding of or engagement in gesturing) that become confounds when completing standardized assessments of receptive language. Thus, assessments that rely on parent report or direct interaction with a test administrator may not fully capture the abilities and underlying knowledge of these children. Visual attention based measures of receptive language reduce task demands placed on participants and have been shown to produce a more accurate measure of receptive vocabulary than parent report measures. However, current gaze-based measures employ a visual paired comparison method. When there are only two items, the similarity of the target to the distractor can have a significant impact on the interpretation of task performance.

The current study evaluates the feasibility of expanding these looking paradigms to include an eight-item array. This Visual Array Task (VAT) combines the theoretical framework of the Intermodal Preferential Looking Paradigm and Looking-While-Listening methods of receptive language with the sequential touching paradigm of object categorization. The use of a larger array of items and the inclusion of a superordinate category contrast could provide a more sensitive measure of receptive language as well as a better understanding of the extent to which early word comprehension reflects knowledge of broader categories.

Results indicate that the tested VAT was both a sensitive measure of receptive vocabulary as well as capable of reflecting gains in category knowledge. This paradigm provides researchers with an inexpensive and efficient task to measure receptive language as well as other general constructs such as category knowledge while reducing behavioral demands placed on the participant. Data here validate the feasibly of using the task to measure receptive vocabulary and category knowledge in children at genetic risk for ASD. Without modification, the VAT can accommodate the testing of minimally verbal to non-verbal populations, as well as populations, such as children with anxiety, that may encounter many of the interaction-based hurdles similar to children with ASD. Thus, the VAT is measure that can be administered uniformly across a broad spectrum of populations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by persistent deficits in social communication and social interaction as well as restricted, repetitive patterns of behavior, interests, or activities (American Psychiatric Association, 2013). The CDC estimates that approximately 1 in 59 children, aged 8, are diagnosed with ASD in the United States (Baio et al., 2018). Although ASD has been conceptualized as a genetically based disorder, present at birth, the earliest reliable age at which a diagnosis can be made is 18 months (Zwaigenbaum et al., 2005). In order to study early emerging social and cognitive processes, researchers have relied on what is known as the infant-sibling paradigm. According to current estimations (Messinger et al., 2015; Ozonoff et al., 2007), approximately 20% of children with an older sibling on the autism spectrum go on to later receive a diagnosis themselves. Knowledge of this familial-based risk has allowed the field to prospectively compare infants at high-risk (HR), younger siblings of children with ASD, to infants at low-risk (LR), younger siblings of typically developing children, of developing ASD. Of course, a recurrence rate of close to 20% means that the majority of HR infants will not go on to receive an ASD diagnosis. However, a subset of these HR children (HR: Non-ASD) experience atypical outcomes such as developmental delay and elevated levels of ASD symptomology (Charman et al., 2017; Messinger et al., 2013).

In addition to the core deficits listed in the diagnostic criteria, individuals with ASD demonstrate differential abilities and/or processing strategies in a number of cognitive domains.

Two areas that have been noted as atypical in children and adults with ASD are receptive language (for review, see Tager-Flusberg et al., 2005) and object category knowledge (for review, see Gastgeb & Strauss, 2012).

Early signs of receptive language and category knowledge emerge within the first year. A preference for human speech helps neonates orient to the social world (Vouloumanos & Werker, 2007) and a sensitivity to categories at two months of age aids in the organization of perceptual experiences (Quinn & Johnson, 2000). Together, these two processes lay the foundation for future language comprehension and production. By 6 months of age, infants have already begun to understand object labels although they will not be able to verbally produce them for several more months (Bergelson & Swingley, 2012; Bergelson & Swingley, 2015). Early deficits experienced in these critical areas of language and category learning may have cascading effects for later development. An exploration of the emergent relationship between receptive language and object category knowledge may thus provide insights into the early cognitive mechanisms driving differences in these domains for individuals with ASD. The present study proposes an examination of the relationship between receptive language and object category knowledge in infants at high familial risk of developing ASD through the use of a novel gaze-based task.

1.1 RECEPTIVE LANGUAGE IN ASD

Impairments and/or delays in language are among the earliest, albeit not uniquely predictive, developmental signs of ASD (Luyster at al., 2011). Delays in language production (expressive language), such as onset of first spoken word, are particularly salient for parents of children later diagnosed with ASD and tend to be the first developmental concern reported to professionals

(Coonrod & Stone, 2004). Early abnormalities in language comprehension (receptive language), such as response to spoken language, are also strong indicators of autism in young children (Lord, 1995). Although no longer included in the diagnostic criteria for ASD, aberrant language profiles remain prevalent (Tager-Flusberg, 2016).

Receptive language is the ability to understand the meaning of spoken words. It is a necessary precursor to spoken language, and shows relative growth over production in typical populations (Fenson et al., 1994). This relationship appears rather intuitive, as one would expect that individuals must learn the meaning of a word before properly producing it. However, a number of studies have identified the reverse relationship, a relative expressive competency advantage, on some measures of language for children with ASD (Charman et al., 2003; Luyster et al., 2007; Luyster et al., 2008; Weismer et al., 2010). Interestingly, the studies conducted by Luyster and colleagues (2008) and Weismer and colleagues (2010) measured more advanced expressive than receptive language in their populations using the *Mullen Scales of Early Learning* (MSEL; Mullen, 1995) and the *McArthur-Bates Communicative Development Inventory* (MB-CDI; Fenson et al., 1996) but more advanced receptive than expressive language using the *Vineland Adaptive Behavior Scales* (VABS; Sparrow et al., 1984). These inconsistencies in competency measured concurrently raise important questions about the validity of the use of standardized measures of language proficiency in ASD populations. It is possible that these assessments may not be fully capturing the early language profiles of children with ASD.

Studies of high-risk infant siblings have also revealed early differences (prior to age 2 years) in receptive language for children that later went on to receive an ASD diagnosis (HR-ASD). Delays in receptive language for HR-ASD infants compared to low-risk infants were reported at 12 months (Mitchell et al. 2006; Zwaigenbaum et al., 2005), and 18 months of age (Mitchell et al., 2006) as measured by the MB-CDI. Delays in receptive language for HR-ASD infants compared to low-risk infants were also reported at 12 months (Zwaigenbaum et al., 2005) and 14 months (Landa & Garrett-Mayer, 2006) as measured by the MSEL.

All of the above report on data collected through the use of standardized assessments that rely on parent report or direct interaction with a test administrator. The most common critique of parental report based measures is the opportunity for reporter bias. This bias becomes of particular concern when parents are asked to report on language comprehension (Tomasello & Mervis, 1994). Not only do parents differ in their abilities to infer language comprehension from their infant or toddler's behavior, but also likely differ in their criteria of what it means for their child to recognize versus understand the referential meaning of a word. This may be particularly true for parents reporting on the younger siblings of children diagnosed with ASD. It is possible that the atypical language profiles exhibited by the older child may skew parent perception of language norms. This bias also appears to differ across cultures. Differential rates of underestimation (Hamilton et al., 2000) and overestimation (Tomasello & Mervis, 1994) of children's receptive vocabularies have been observed between British and North American parents using equivalent versions of the CDI.

To reduce such bias and variability, researchers have turned to experimenter-administered measures of infant and toddler competencies such as the Bayley Scales of Infant Development (Bayley, 1993) and the MSEL. However, previous studies have shown that infants and toddlers with ASD may have language, social, and communicative deficits (i.e. gaze aversion and reduced understanding of or engagement in gesturing) that become confounds when completing standardized assessments (Kasari et al., 2013; Brucker et al., 2007; Plesa Skewer et al., 2015). For example, certain individuals with ASD may exhibit characteristics such as orientation deficits, social communication deficits, and restricted object use that make it challenging to complete such assessments (Bruckner et al., 2007). Other difficulties for individuals with ASD during interactive assessments include: difficulties understanding the pragmatics of the instructions, difficulties with social responsiveness to the examiner, and being unable to point or produce similar gestural replies (Tager-Flusberg, 1999). Thus, standardized assessments may not fully capture the abilities and underlying knowledge of infants at risk for and children with ASD.

Two less behaviorally demanding measures of receptive language that significantly limit the above concerns are the Intermodal Preferential Looking Paradigm (IPLP; Golinkoff et al., 1987) and the looking-while-listening task (LWL; Fernald et al., 1998; Fernald et al., 2008). Based on the visual paired comparison method (Fantz, 1958), the IPLP presents children with two visual stimuli while a verbal label matching one of the items is simultaneously played. For example, a child could be presented with an image of an apple and an orange while hearing the word "apple" spoken. If infants preferentially fixate the item that is being labeled, it is inferred that the child has linguistic knowledge of the stimulus. In this case, if the child looked to the apple longer than the orange when hearing the word "apple", it would suggest that the child associates the spoken word label with the object. Thus, the IPLP provides a behavioral measure of receptive language without relying on parent report, examiner interaction and/or, as is the case for the MSEL, more demanding verbal or gestural responses from the participant.

The LWL procedure was developed as an extension of the IPLP. Therefore, the same basic task principles apply. Children are presented with two images and a single verbal label. However, instead of relying on overall preference of looking to one picture over the other, the

LWL procedure codes moment-by-moment eye movements to determine changes in gaze related to the onset of the linguistic stimulus. For example, in our previous scenario, experimenters would measure where the child was looking during relevant points in the linguistic signal, such as phonemic onset "**æ**", which would quickly distinguish the two words from one another. Essentially, the LWL task provides a moment-by-moment measure of visual attention during speech processing.

Both the IPLP and the LWL procedure have demonstrated the ability to capture growth in receptive vocabulary across age for typically developing infants and children (Golinkoff et al., 2013; Fernald et al., 2008). In fact, these methods may provide a more accurate measure of receptive vocabulary than parent report measures (Houston-Price et al., 2007). Using the IPLP Houston-Price and colleagues (2007) were able to verify infant knowledge of words parents had previously reported as "unknown" suggesting a parental underestimation of infant receptive vocabularies.

Research findings also verify the ability of both the IPLP and the LWL task to identify differences in receptive vocabulary between children and adolescents with ASD and typical controls (Bavin et al., 2014; Brady et al., 2014; Skerwer et al., 2016; Venker et al., 2013). However, only one study to date has examined the feasibility of using these procedures in infant-sibling populations. Chita-Tegmark and colleagues (2015) used the LWL procedure to test for differences between receptive vocabularies in high- and low-risk infants at 18, 24, and 36 months of age. Gaze results for target fixation accuracy identified lower performance for the HR group at 36 but not 18 or 24 months of age. There were no differences found between groups for their measure of reaction time (speed to target fixation following the onset of the verbal label). The authors conclude that this indicates intact language processing efficiency for children at risk for

ASD but a potential difference in word acquisition abilities leading to a more restricted receptive vocabulary at 36 months of age.

1.2 CATEGORIZATION IN ASD

The first studies examining categorization abilities of individuals with ASD suggested a lack of deficit in this domain (Tager-Flusberg, 1985a, 1985b; Ungerer & Sigman, 1987). However, these studies used relatively simple tasks (e.g. sorting, and matching-to-sample) that relied on easily identifiable perceptual features. More recent work has focused on the processes (e.g. prototype abstraction) that underlie the categorization of more complex stimuli including natural categories, faces, and complex dot patterns. These studies have identified a number of differences between individuals with ASD and typical controls (Gastgeb & Strauss, 2012; Gastgeb, Strauss, & Minshew, 2006; Gastgeb et al., 2012). Combined, these data suggest that there may be two underlying differences in the way that individuals with ASD process category information. First, while individuals with ASD are no different with very simple categories that have criterial features (such as common color or shape), they demonstrate difficulties with more complex categories (e.g. faces). These complex categories are more similar to natural categories that do not have clear boundaries, definitive features, and vary in typicality (Mervis & Rosch, 1981; Rosch, 1978). Second, individuals with ASD may process these more complex or natural categories differently and have difficulties forming central or prototypical representations of the categories (Gastgeb & Strauss, 2012).

Various methods have been developed to study category knowledge in infants and young toddlers (Mareschal & Quinn, 2001). One task that has been used with toddlers is the sequential touching procedure (Mandler et al., 1987; Ricciuti, 1965; Starkey, 1981). In the sequential touching procedure infants are presented with a group of objects that form two or more distinct categories (e.g. four animals and four vehicles). Children are encouraged by an experimenter to manipulate the objects and then given several minutes in which they can interact and play with the objects in any way they choose. These interactions are recorded and the number and sequence of item touches are scored after the play session. Interestingly, children naturally begin to touch the objects in sequences that indicate their awareness of category relationships. When children make more sequential touches to objects of the same category than would be expected by chance, category knowledge is inferred. Various types of category contrasts can be used depending on the specific aims of the study. A common question examined by this paradigm is whether young children have knowledge of superordinate (global; Mandler et al., 1987), basic, and subordinate classes of objects. Findings indicate that between 18 and 30 months of age, infants spontaneously categorize based on tests of superordinate level contrasts only. At 30 months of age toddlers begin to spontaneously categorize based on tests of basic level contrasts as well (Mandler et al., 1988; Mandler et al., 1991).

One published study (Ungerer & Sigman, 1987) and one unpublished dissertation (Vitrano, 2015) have explored categorization behavior using sequential touching tasks for children with ASD. Ungerer and Sigman (1987) observed sequential touching behaviors in 16 children, aged 3 to 6 years, diagnosed with ASD using three different contrast types (all superordinate) to assess knowledge of form, color, and function. Stimuli for form and color contrasts were wooden square, triangle, and circle shapes. Stimuli for function contrasts were small miniatures representing animals, fruits, vehicles, and furniture. No differences in categorization based on form, function, or color were observed between individuals with ASD and the IQ-matched typical or “mentally retarded” controls. This lack of participant difference may have been due to the use of simple criterial categories (i.e. shape and color) as describe above.

In his dissertation, Vitrano examined sequential touching behaviors for 10 children, also aged 3 to 6 years, diagnosed with ASD during three tasks: a superordinate level contrast, a basic level contrast, and a subordinate level contrast. The children demonstrated evidence of categorization for the superordinate and basic level contrasts but not the subordinate level contrast. These findings are in agreement with those reported by studies examining object category formation in typically developing children. Unfortunately, this study did not include data from a typically developing control group. Therefore, it is unknown if the performance of the children with ASD would have directly differed from that of typically developing children.

1.3 INTERSECTION OF LANGUAGE AND CATEGORIZATION

While there is some debate as to the direction of the relationship between categorization and language development, Waxman and Markow (1995) offered the first evidence supporting a link between the two in infancy. Examining the noun-category linkage in early lexical development Waxman and Markow found that for 12- to 13-month-old infants object category labels learned during a familiarization phase would generalize to new exemplars during a testing phase. Findings reported since that seminal study support the notion that infants are able to generalize object labels from an exemplar to similar objects belonging to the same category based on perceptual, functional, and conceptual properties (for a review, see Ferguson & Waxman, 2017). For example, when a dog is labeled for infants, infants know that the label *dog* does not only apply to that one specific dog, say a family pet, but also to other dogs they may encounter. Based on the common properties shared between dogs, like fur, overall shape, barking, and independent motion, infants place the exemplar within the basic category of dog or even the global category of animals. Of course at some level this must be a bidirectional process. While categories can help infants learn object names, labeling objects can also aid in the recognition of commonalities between object category members. For example, naming basic level objects, such as a dog or a cat, with the superordinate label animal promotes attention to the features those two objects have in common- like fur and independent motion, but not barking.

The two studies measuring sequential touching in ASD have also measured the link between receptive language and object category knowledge in ASD (Ungerer & Sigman1987; Vitrano, 2015). First, although Ungerer and Sigman (1987) did not find significant group differences in categorization for children with ASD, they did observe a relatively weakened link between measures of object categorization and receptive language for their children with ASD. Conversely, Vitrano (2015) did not observe a significant correlation between performance on the sequential touching task and receptive language as measured by the Peabody Picture Vocabulary Test (PPVT-II; Dunn & Dunn, 1997).

1.4 THE PRESENT STUDY

Language comprehension is a difficult construct to measure. Assessments based on parent report rely on the parents to infer language and vocabulary knowledge through observations of their children. There may be large differences in how accurately parents report this information. Moreover, standardized laboratory based developmental assessments place high task demands on children to engage in pre-arranged tasks such as pointing to photos in a book or following a set of verbal task instructions (e.g. “give the doll to mommy”). This is problematic as these measures of receptive language are also dependent on successful interaction with a (often unfamiliar) test administrator as well as gestural behaviors. These added restrictions may become particularly challenging for populations that are at risk for language delay and communication deficits such as children with ASD. While visual attention based measures of receptive language such as the IPLP and the LWL task are less behaviorally demanding, they also have limitations.

A significant concern of the IPLP and LWL task is that they require participants to distinguish a verbally labeled target from a single alternative “distractor” item. According to the mutual exclusivity (Markman & Wachtel, 1988) and novel name-nameless category (N3C; Golinkoff et al., 1994) principles, children may preferentially fixate the target object not because they know the label of the target, but rather, because they know the label of the distractor. When presented with two objects, one familiar and one novel, children will attribute a novel object label to the unknown category. This ability to fast-map is often linked to the vocabulary growth spurt that occurs around 18 months of age (Mervis & Bertrand, 1994).

When there are only two items, the similarity of the target to the distractor can also have a significant impact on the interpretation of task performance (Arias-Trejo & Plunkett, 2010). For example, imagine a scenario in which a child is presented with two images and told to “look at the dog”. If that child is presented with a dog vs. cow, they may not look preferentially to the dog because of the perceptual similarity of the two items. Here, a lack of looking would be interpreted as the child not yet having acquired the object name for dog. In contrast, the same child might look preferentially to the dog when it is paired with a car or even another animal that is perceptually very different (e.g. a turtle). In this case, the child’s behavior would be (likely incorrectly) interpreted as knowledge of the object label for dog. Thus, results are highly dependent on the nature of any given comparison.

The present study addresses this limitation by altering the methodology of the IPLP to include a multiple object array. By using eye tracking procedures it is possible to present children with more than just two items. Previous work by Brady et al. (2014) has successfully expanded this task to a four-object array for children ages 42-82 months (approximately 3 ½ - 7 years of age). Despite the increased number of distractor objects, both typically developing children and children with ASD were able to reliably identify the named target object.

The current study looks to evaluate the feasibility of expanding this looking paradigm to include an eight-item array. The expansion of the task to include eight items mimics the object manipulation paradigm. Therefore, the task will also include a superordinate category contrast. Four objects will belong to one superordinate category (e.g. animals) and four objects will belong to another superordinate category (e.g. items of clothing). This may afford the measurement of category knowledge through not only overall looking preference but also the scoring of sequential looks. Such a task would be well suited to the study of infant and ASD populations, as it requires minimal behavioral and no verbal responses from the participants.

This visual array task (VAT) combines the theoretical frameworks of the IPLP and LWL paradigms of receptive language with the sequential touching paradigm of object categorization. The use of a larger array of items and the inclusion of a superordinate category contrast could provide a more sensitive measure of receptive language as well as a better understanding of the extent to which early word comprehension reflects knowledge of broader categories. It also allows for the examination of the relationship between receptive language and object category knowledge as measured by the same behavioral task. Through the use of this novel gaze-based task and the infant sibling paradigm, the present study aims to:

1) Validate the visual array task (VAT) as an online measure of receptive vocabulary and category knowledge in a typically developing population.
2) Assess group differences in receptive vocabulary and category knowledge using the VAT based on (a) genetic risk for ASD and (b) later outcome classification.

1.4.1 Predictions

Typically developing toddlers. The inclusion of seven distractor items will likely lead to an increase in task difficulty; however, this added complexity is not expected to disrupt target identification for object labels that are commonly known by 16- and 24-month-old toddlers. Based on findings from the IPLP and LWL paradigms, it is predicted that toddlers will understand the object label and fixate the target item for a longer duration and more frequently than would be expected by chance. Based on the expected growth in receptive language from 16 to 24 months of age, it is predicted that toddlers will fixate the target object longer at 24 months of age compared to 16 months of age. Both the IPLP and the LWL task have demonstrated a degree of convergent validity with standardized measures of receptive language (Golinkoff et al., 2013; Law & Roy, 2008). Therefore, it is predicted that this expansion of the IPLP will as well.

As this is the first task to probe category knowledge in such a way, there are no specific predictions related to indictors of category knowledge. However, we did decide to center analysis pertaining to this question around the other-category objects. Given that the audio recording is used to orient the participants to the target object, it was unclear what the elevated status of the target meant for the remainder of the target category members.

Toddlers at high- and low-risk for ASD. Based on previous research citing early deficits in receptive language for infants and toddlers with ASD, it is predicted that there will be a main effect of risk indicative of greater target identification by the LR infants. While we predict that both groups will fixate the target item above chance, it is believed that the added complexity of the task will induce risk group differences not observed by Chita-Tegmark and colleagues (2015) at these ages. As there are no prior studies on object category knowledge in toddlers at risk for ASD, the tests for differences between the HR and LR groups on measures of category understanding are exploratory in nature.

2.0 METHOD

2.1 PARTICIPANTS

All participants were recruited by the Autism Center of Excellence (ACE) at the University of Pittsburgh and drawn from a larger study conducted by the Center for Infant and Toddler Development (ITDC). The sample consists of infant siblings of children with ASD (high-risk infants; HR) and infant siblings of typically developing children (low-risk siblings; LR). HR participants had at least one older sibling with an ASD diagnosis confirmed by the Autism Diagnostic Observation Schedule-Generic (ADOS-WPS; Lord et al., 2001) and the Autism Diagnostic Interview-Revised (ADI-R; Lord et al., 1994). LR participants had at least one typically developing older sibling as well as no first or second-degree relatives with ASD. Additional exclusion criteria for all participants were a birth weight less than 2500 grams, problems with pregnancy, labor or delivery, traumatic brain injury, prenatal illicit drug or alcohol use, and/or birth defects.

Infants were recruited into the broader study at 6 months of age and followed longitudinally to 4 years of age. Participants were initially seen at the ITDC at 6, 11, and 16 months of age to participate in a number of eye tracking tasks as well as developmental assessments including the Mullen Scales of Early Learning (MSEL; Mullen, 1995) and the MacArthur-Bates Communicative Development Inventories (MB-CDIs; Fenson et al., 1993).

Participants were asked to return at 24, 36, and 48 months of age for follow-up and developmental evaluations. During these evaluative visits, parents again completed a MB-CDI and children were assessed by the MSEL. Additionally, all children received an ASD evaluation through administration of the ADOS-G.

Based on the above assessments and clinical judgment, participants were categorized into one of three possible outcome groups: ASD, non-typically developing (NT), or typically developing (TD). The categorization of each child into one or more of these outcome groups was reviewed and confirmed by clinical opinion of a licensed psychologist. The criteria for inclusion into each outcome group were as follows:

1) *ASD*: Children in the ASD group met spectrum cut-offs on all three ADOS-WPS total scores, which were then reviewed and approved by a clinical psychologist.
2) *NT*: Children in the NT group demonstrated atypical development as indicated by at least one of the following: global developmental delay, language delay, and/or social communication concerns. These were defined as follows:
    a) *Global developmental delay:* Children identified as having global developmental delay scored more than 1.5 standard deviations below the normative mean on the Visual Reception and Receptive Language subscales of the MSEL.
    b) *Language Delay:* Children were identified as having a language delay if they scored more than 1.5 standard deviations below the normative mean on the Receptive Language and/or Expressive Language subscales of the MSEL. Children with a Words Produced score below the 10^{th} percentile on the MB-CDI were also identified as having a language delay.

c) *Social Concerns:* Children with social concerns met at least spectrum cutoffs on only the ADOS-WPS Social Interaction total or scored within 2 points of spectrum cutoffs on the combined Communication and Social Interaction totals.

3) *TD*: Children that did not meet criteria for the ASD or NT groups were categorized as typically developing.

To be included in the analyses of this study, infants had to have completed the visual array task (VAT) at 16 and/or 24 months of age. The study sample includes data from 88 toddlers (51 Males; 10 Racial or Ethnic Minorities). Twenty-three children (9HR, 14LR) completed the VAT at both the 16- and 24-month visit. These 23 children will be referred to as the longitudinal cohort. Demographic information and standardized assessment scores for the longitudinal cohort may be found in Table 1 and Table 2 respectively. An additional 65 children completed the VAT at only one visit: 24 (18 HR, 6 LR) 16-month-olds and 41 (13 HR, 28LR) 24-month-olds. To allow for the greatest power to detect group differences at each age, the cross-sectional sample will include these 65 children as well as the children in the longitudinal cohort. Demographic information and standardized assessment scores for the cross-sectional sample are displayed in Table 3 and Table 4 respectively.

Eighty-four of the 88 toddlers successfully completed the outcome assessment and were classified into outcome groups at 24 (N = 3), 36 (N = 34), and/or 48 (N = 47) months of age. In the event that a participant was assessed at multiple visits, the latest assessment was used for classification. For the longitudinal group, outcome data was recorded for all 23 children. Of these 23, 1 child was diagnosed with ASD, 3 children were classified as NT, and 19 were classified as TD. In the cross-sectional 16-month group, 4 children were diagnosed with ASD, 3 children were classified as NT, and 14 were classified as TD. In the cross-sectional 24-month group, 4 children were diagnosed with ASD, 4 children were classified as NT, and 32 were classified as TD. Demographic information and standardized scores for the outcomes groups are listed in Tables 5-6 (longitudinal cohort) and Tables 7-8 (cross-sectional sample).

2.2 STIMULI

Toddlers were shown 12 stimulus trials over the course of the testing procedure. Each trial contained an array of eight prototypical color illustrations of common objects (see Figure 1). These illustrations were approximately 7.7 x 7.7 degrees of visual angle each. Importantly, all objects were chosen to represent vocabulary words that 16- and 24-month-old toddlers would be familiar with (as determined by normative scores of the MB-CDI and MSEL). The objects presented were also chosen from six superordinate categories that children of this age are familiar with: vehicles, clothing, animals, food, furniture, and utensils (Mandler et al., 1991; Ross, 1980). The object members of each superordinate category are listed in Table 9. Each trial (i.e., picture) included a superordinate contrast. That is, four objects from each array belonged to one superordinate category (e.g. vehicles) while the remaining four items belonged to another superordinate category (e.g. animals). The locations of objects on the screen were randomized using a 5 (width) x 4 (height) grid system. This ensured that children did not make anticipatory looks to any one area of the screen while completing the task. The target item appeared in the six central grid locations on approximately 20% of the trials and in the 14 peripheral grid locations on approximately 80% of the trials (see Figure 2).

For each trial, toddlers also heard an audio recording of a female voice naming one of the objects on the screen. This object is known as the “target”. Once the array was projected onto the screen, the voice began “(Target). Look at the (target). Where is the (target)? (Target). See the (target)?” This particular phrasing was chosen so that the toddlers heard the target name said both as an isolated word and embedded within a sentence. This was done to help to maximize their ability to spontaneously look at the labeled object. While the phrasing remained the same throughout the testing session, the target object changed with each trial. For a trial in which the target item was the image of a dog, participants heard “Dog. Look at the dog. Where is the dog? Dog. See the dog?”

2.3 APPARATUS

The eye tracking portion of the task took place in a dark, quiet room that resembled a small movie theater. Each participant was seated in a highchair approximately 152 cm from a large projection screen (69 x 91 cm) with their guardian seated next to them. Guardians were instructed not to point to or talk about the images projected onto the screen, but were encouraged to comfort the toddlers if necessary. Eye movement data was recorded by a standalone Tobii X120 eye tracker positioned on a table in front of the participant, approximately 81 cm from the screen. Using Tobii Studio software (Version 2.0.6), the stimuli were rear-projected on the screen and the participant’s eye movements were recorded at a sampling rate of 60 Hz, accuracy of 0.5 degrees of visual angle, spatial resolution of 0.2 degrees, and drift of 0.3 degrees. Raw eye movement data was then converted into fixations using the Tobii fixation filter (Olsson, 2007).

2.4 PROCEDURE

2.4.1 Visual array task (VAT)

Once the toddler and guardian were seated comfortably, a cartoon was played to orient the child toward the screen and maintain his or her attention. When the toddler was quiet and attending to the screen, the cartoon was turned off and a calibration period began. During calibration, toddlers were visually prompted by a moving target to orient their gaze to a total of five predetermined locations on the screen: the center of the screen and each of the four corners. These targets were small, brightly colored objects that simultaneously produced a slight motion (jitter or oscillation) and corresponding sound. Once the experimenter determined that the toddler was attending to the current location, they manually advanced the target to the next location. If the Tobii eye tracker and software system detected both the right and left eye of the toddler at each calibration target location, the calibration was considered successful. This process was repeated until a successful calibration was obtained.

Following a successful calibration, toddlers viewed 12 stimulus presentation trials that lasted approximately 10 seconds each. In between each trial, a short cartoon was played to reorient the child's gaze to the center of the screen. If the participant became upset or distracted during the task, stimulus presentation was halted until the child was calm and attending to the screen again. In total, each toddler saw 12 different object arrays and heard 12 different target words. The presentation of all possible target objects (24) and six superordinate category contrast combinations (out of a possible 15) was counterbalanced across participants. This resulted in six distinct stimulus sets (see Figure 3).

2.4.2 Standardized measure of language comprehension

The *Mullen Scales of Early Learning* (MSEL; Mullen, 1995) was administered for both the 16- and 24-month age groups. The MSEL is a standardized, observation based assessment designed to measure cognitive development in children from birth to 68 months of age. MSEL administration took place in a small quiet room that minimized distractions for the toddlers. Each session was conducted by a trained member of the research team and video recorded to allow for secondary score verification and clinical review. In rare cases where children became fatigued or upset, the assessment was carried out over two visits to the laboratory. The four cognitive subscales (visual reception, fine motor, receptive language, and expressive language) were combined to create an Early Learning Composite to be used as a measure of overall cognitive functioning. The score from the receptive language scale was also considered independently as a measure of language comprehension.

The *MacArthur-Bates Communicative Inventories* (CDIs; Fenson et al. 1993) were administered for both the 16- and 24-month age groups. The MB-CDIs are a parent report measure of global language and communicative development in infants from 8 to 18 months of age (Words and Gestures, WG) and toddlers 16 to 30 months of age (Words and Sentences, WS). Guardians of toddlers in the 16-month age group completed the MB-CDI: WG form while parents of toddlers in the 24-month age group completed the MB-CDI: WS form. A trained member of the research team described the protocol for form completion to the participant's guardians during the consenting procedures. Any questions related to the assessment's completion were answered during the study visit. Guardians then completed the forms at home and returned them during a subsequent visit to the ITDC or through the mail via preaddressed and stamped envelopes. Developmental percentiles from the MB-CDIs (WG: phrases understood, words understood, words produced; WS: words produced, grammatical complexity, mean of the three longest utterances) were used as an overall measure of language and communication competency. The MB-CDI does not have a composite score associated with it by design.

2.5 DATA REDUCTION

Trials in which toddlers failed to fixate at least one of the eight objects were eliminated from analyses (38/1,332; approximately 3%). The number of trials eliminated did not vary by age or risk group.

2.5.1 Areas of interest (AOIs)

The number and duration of fixations (stabilized gaze on a single location) made to the objects was determined by creating AOIs for each of the eight objects in the arrays. Using Tobii Studio software, a square AOI was drawn around each object that is approximately the size of the object (190 x 190 pixels). Fixations to these AOIs were then classified into three types. Fixations to the object that was named were defined as target fixations (1 object per trial). Fixations to the 3 objects belonging to the same superordinate category of the target were defined as target category fixations (3 objects per trial). Fixations to the objects that did not belong to the superordinate category of the target were defined as other-category fixations (4 objects per trial; see Figure 4).

2.5.2 Measures of visual attention distribution

To determine how toddlers distributed attention among the objects on the screen, the following four gaze-based variables were calculated. Proportions were used to control for variation in total looking time to the stimuli among participants.

1) *Proportion of Target Fixations:* The total number of fixations to the target divided by the number of fixations to all objects- summed across the 12 trials
2) *Proportion of Target Fixation Duration:* Total time spent fixating the target divided by the total time spent fixating all the objects- summed across the 12 trials
3) *Proportion of Other-Category Fixations:* The total number of fixations to the four objects belonging to the other-category (the non-target category) divided by the total number of fixations to all objects- summed across the 12 trials
4) *Proportion of Other-Category Fixation Duration:* The total time spent fixating the four objects belonging to the other-category (the non-target category) divided by the total number of fixations to all objects- summed across the 12 trials

2.5.3 Measures of systematic scanning

In traditional tasks of sequential touching and object manipulation contact of an object by hand or with an object is considered a touch. A successive touch of the same object does not count toward the run length (e.g. picking up an object, placing it down, and then picking up the same object up again). This behavior would be considered analogous to making two or more sequential fixations within the same AOI. Therefore, each visit to one of the eight object AOIs was considered a “touch”. Using a video playback of the eye tracking session, trained coders manually scored the number and sequence of visits between the eight objects. Each time a toddler’s gaze entered and exited an AOI was scored as a visit. A single visit may contain multiple fixations (see Figure 5 for a visual representation).

The visit sequences were then used to create two measures of systematic scanning. The first measure is the number of runs toddlers make throughout the testing session. A run will constitute two or more sequential visits to objects belonging to a single superordinate category. The second measure is run length of successive looks to the four members of each category. It will be a proxy for whether or not toddlers are scanning the objects systematically (reflecting an impact of category) or randomly.

1) *Number of Runs (NRuns):*
    a. *NRuns Target Category*: The number of runs made between objects belonging to the superordinate category of the target object- summed across all 12 trials.
    b. *NRuns Other Category*: The number of runs made between objects belonging to the superordinate category unrelated to the target object- summed across all 12 trials.

2) *Mean Run Length (MRL):*
    a. *MRL Target Category:* The length of all runs within the target category divided by the total number of runs- summed across all 12 trials
    b. *MRL Other-Category:* The length of all runs within the other-category divided by the total number of runs- summed across all 12 trials

3.0 RESULTS

The current study includes a mixture of cross-sectional and longitudinal data. To allow for the greatest power to detect group differences at each age, cross-sectional analyses included data from participants in the longitudinal cohort. Strong correlations between the variables of fixation count and fixation duration were observed; therefore, only results for fixation duration are reported here.

3.1 TYPICALLY DEVELOPING TODDLERS

The first aim of this study was to explore the feasibility of using the VAT as an online measure of receptive vocabulary and object category knowledge in a typically developing sample of toddlers. To this end, the following analyses were carried out using data from the LR-TD toddlers only. If you recall, this sample includes data from 17 toddlers collected at 16 months of age and 37 toddlers collected at 24 months of age. Embedded in this broader cross-sectional sample is also a longitudinal cohort of 13 toddlers. Demographic information and standardized assessment scores for this subsample may be found in Tables 10 -11 (cross-sectional sample) and Tables 12-13 (longitudinal sample). Means and standard deviations for the primary variables of interest may be found in Table 14 (cross-sectional sample) and Table 15 (longitudinal subsample).

Are toddlers able to accurately identify the target object? Of primary concern is whether toddlers will preferentially fixate a target object within a larger eight-item array. To this end, analyses were conducted on the cross-sectional sample to determine if the proportion of target fixation duration was greater than the chance value of 1/8 (0.125). This chance value represents an equal distribution of fixation duration between all eight items in the array. If toddlers fixate the target object for a proportion of time greater than 1/8, a preference for the target object will be inferred. One sample t-tests revealed that typically developing toddlers fixated the target object for a longer duration than would be expected by chance at both 16 (M = 0.18, SD = 0.06; $t(16) = 4.05$, $p = .001$) and 24 months of age (M = 0.28, SD = 0.09; $t(36) = 10.44$, $p < .001$). This indicates that, despite the added level of complexity, children as young as 16 months of age are able to accurately identify a known target item from an eight-item array.

Do toddlers demonstrate reduced attention to other-category items? Children with knowledge of the target object category may preferentially look to members of the target category and/or inhibit looking to members of the other superordinate category. One sample t-tests conducted on the cross-sectional sample revealed that typically developing toddlers fixated the other-category objects for a shorter duration than would be expected by chance (4/8; 0.5) at both 16 (M = 0.44, SD = 0.08; $t(16) = -3.27$, $p < .01$) and 24 months of age (M = 0.35, SD = 0.07; $t(36) = -11.80$, $p < .001$). Again, this chance value of 4/8 (0.5) assumes an equal distribution of visual attention across all eight objects in the array. The findings that toddlers are looking less than chance indicates that in addition to preferentially fixating the target object, toddlers also may inhibit looking to the other-category objects at both 16 and 24 months of age.

Does the Visual Array Task capture developmental change over time? Toddlers in the longitudinal cohort demonstrated approximately a 9% increase in their proportion of target fixation duration between 16 (M = 0.18, SD = 0.05) and 24 (M = 0.27, SD = 0.08) months of age. A paired samples t-test confirmed that between 16 and 24 months these toddlers demonstrated a developmental increase in their proportion of target fixation duration ($t(12) = -4.98$, $p < .001$). This task also measured a corresponding decrease in other-category object fixation across age ($t(12) = 3.94$, $p < .01$). Importantly, there were no observed differences in looking to the three other target-category objects across age ($t(12) = -.24$, $p = .82$). This indicates a reciprocal relationship between target looking and other-category looking. As toddlers begin to preferentially fixate the target item they selectively decrease their fixation of the other-category items (see Figure 6).

Does performance on the VAT correlate with standardized measures of receptive language? To assess the convergent validity between the VAT, MSEL and, the MB-CDI correlation matrices were computed separately for each cross-sectional age group.

The matrix for the 16-month-old group included the following variables: (1) proportion of target fixation duration (2) proportion of other-category fixation duration (3) ELC derived from the MSEL (4) receptive language t-score derived from the MSEL, (5) phrases understood percentile from the MB-CDI Words and Gestures, (6) words understood percentile from the MB-CDI Words and Gestures, and (7) words produced percentile from the MB-CDI Words and Gestures. Results of the Pearson's correlations indicated that there were no significant associations between the variables derived from the VAT and scores from the MSEL or the MB-CDI at 16 months of age (see Table 16).

The matrix for the 24-month-old group included the following variables: (1) proportion of target fixation duration (2) proportion of other-category fixation duration, (3) ELC derived from the MSEL (4) receptive language t-score from the MSEL, (5) words produced percentile from the MB-CDI Words and Sentences, (6) grammatical complexity percentile from the MB-CDI Words and Sentences, and (7) mean of the three longest utterances percentile from the MB-CDI Words and Sentences. Results of the Pearson's correlations indicated that there were no significant associations between the proportion of target fixation and the variables derived from the MSEL or the MB-CDI at 24 months of age (see Table 17). However, there were a number of significant negative associations between the proportion of other-category fixation duration and the other measures. First, there was a significant negative association between the proportion of fixation duration to the target and the proportion of fixation duration to the other-category objects ($r(35) = -0.82$, $p < .001$). This finding confirms the observed reciprocal relationship wherein as toddlers look more to the target object they reduce looking to the other-category objects. Second, moderate negative associations were observed between the proportion of looking to other-category objects and the ELC standard score ($r(35) = -0.45$, $p < .001$) as well as the receptive language t-score ($r(35) = -0.45$, $p < .001$). This indicates that increased looking to the other-category objects is associated with lower overall performance scores on the MSEL as well as lower receptive language subscale scores in particular.

Does performance on variables derived from the VAT at 16 months of age correlate with performance at 24 months of age? To assess how well the experimental measure tracks receptive language growth relative to the standardized measures of the MSEL and MB-CDI a correlation matrix including the same variables listed above was calculated using data from the longitudinal cohort. A significant positive correlation was observed for the VAT task between the proportion of target fixation duration measure at 16 months of age and 24 months of age ($r(11) = .61$, $p = 0.03$). Pearson's r correlations for the subscales of the MB-CDI ranged from $r(11) = .66$ to $r(11) = .93$ (see Table 18). This indicates a moderate to very strong relationship between the scores collected at 16 and 24 months of age for all subscales of the measure. A significant correlation was not observed between scores collected at 16 and 24 months of age for the receptive language subscale of the MSEL ($r(10) = .38$, $p = 0.23$) or the ELC ($r(10) = .42$, $p = 0.17$).

Do toddlers demonstrate sequential looking during the Visual Array Task? Sequential touching during object manipulation tasks is a natural behavior children engage in. When presented with an array of objects belonging to two or more categories children begin to touch objects in a sequence that reflects understanding of category relations (Rosch, 1978). It is unknown whether a similar tendency toward sequential looking within categories will emerge when children are presented with a visual array of objects. Collapsing across age, toddlers made on average 18 runs within the target category (MRL target = 3.58) and 13 runs within the other-category (MRL other-category = 3.03) throughout the 12 trials of the VAT. This suggests that toddlers do engage in sequential looking when presented with a visual object array.

Do patterns of sequential looking change across development? Paired samples t-tests were calculated for NRuns target category, NRuns other-category, MRL target category, and MRL other-category to determine if toddlers in the longitudinal cohort demonstrated an increase in the tendency to engage in sequential looking between 16 and 24 months of age. Results indicated that between 16 and 24 months of age toddlers increase their tendency to make runs within the target category ($t(12) = -3.99$, $p < .01$) and decreased the length of runs they make within the other-category ($t(12) = 3.57$, $p < .01$).

Do sequential looking behaviors indicate underlying category knowledge? In order to test the impact that category membership may have had on scan patterns two ratio variables were created. The first is the number of runs ratio (NRuns ratio) calculated as the number of runs made within the target category divided by the number of runs made within the other-category. The second is the mean run length ratio (MRL ratio) calculated as the average run length made within the target category (MRL Target Category) divided by the average run length made within the other-category (MRL Other Category). An NRuns ratio value greater than one indicates more frequent sequential looking within the target category. A MRL ratio value greater than one indicates longer sequential looking within the target category. To determine if toddlers demonstrated a preference for sequential looking within the target category compared to the other category, one-sample t-tests comparing the NRuns ratio and MRL ratio to the no difference value of 1 were calculated for each cross-sectional age group. Results indicated a preference for making runs within the target category (NRuns ratio; $t(16) = 2.22$, $p = .04$) but no difference in the length of runs between the target and other category (MRL ratio; $t(16) = 0.42$, $p = .68$) for the 16-month group. The 24-month group demonstrated more frequent ($t(36) = 6.61$, $p < .001$) and longer runs within the target category ($t(36) = 7.06$, $p < .001$) compared to the other category.

Do measures of receptive vocabulary correlate with measures of category knowledge? To assess the relationship between measures of receptive language and object category knowledge during a single testing session, correlation matrixes were computed separately for toddlers in the cross-sectional sample at 16 and 24 months of age. The eye tracking variables included in the matrix were: (1) proportion of fixation duration target, (2) proportion of fixation duration other-category, (3) NRuns ratio, and (4) MRL ratio (see Table 19). No significant associations between the measure of receptive vocabulary (proportion of target fixation) and the three measures of category knowledge were observed for the 16-month-old group.

However, significant associations between the measure of receptive vocabulary and all three measures of category knowledge were observed for the 24-month-old group. A strong negative association was observed between the proportion of target fixation duration and the proportion of other-category fixation duration ($r(35) = -.82$, $p < .001$). Again, this indicates a reciprocal relationship between time spent looking to the target object and time spent looking to the four other-category objects. A strong positive association was observed between the proportion of fixation duration to the target item and the NRuns ratio ($r(35) = .70$, $p < .001$). This finding indicates that as toddlers increase time spent looking to the target item they are also increasing the proportion of runs that they make within the target category. Finally, a moderate to low positive association between the proportion of target fixation duration and the MRL ratio was observed ($r(35) = .33$, $p = .05$). This indicates that as toddlers increase the proportion of time spent looking to the target object, they also increase the length of runs they make between objects belonging to the target category compared to items belonging to the other superordinate category.

3.1.1 Summary of typical findings

The results reported above validate the use of the VAT in typical populations. Despite added difficulty, children at 16 and 24 months of age were able to reliably identify the target object. Performance on this task was also shown to measure an underlying knowledge of superordinate categories. As toddlers increased their looking to the target object, they demonstrated a reciprocal decrease in looking to other-category objects. While target fixation did not correlate with other standard measures of early language comprehension, there was a strong correlation between performance on this task at 16 and 24 months of age. Results indicated that there were also differential impacts of target label and category knowledge on sequential looking patterns across age groups. Between 16 and 24 months of age toddlers increased the number of runs made within the target category and decreased the length of runs made within the other-category. At 16 months of age, a preference to engage in sequential looking between members of the target category was reflected in the number of runs made within the target category. By 24 months of age this preference was reflected in both the number of runs as well as the length of runs made within the target category. This suggests that an increase in target category knowledge between 16 and 24 months of age results in greater sequential fixation of objects belonging to the target's superordinate category.

3.2 TODDLERS AT HIGH- AND LOW-RISK FOR ASD

The second aim of this study was to assess any differences in performance on the VAT that may arise between toddlers at high- and low- genetic risk for ASD. If you recall, this sample includes data from 47 toddlers collected at 16 months of age (N_{LR} = 20 and N_{HR} = 27) and data from 64 toddlers collected at 24 months of age (N_{LR} = 42 and N_{HR} = 22). Embedded in this broader cross-sectional sample is also a longitudinal cohort of 23 toddlers (N_{LR} = 14 and N_{HR} = 9). Means and standard deviations for the primary variables of interest may be found in Table 20 (cross-sectional sample) and Table 21 (longitudinal subsample).

Do HR and LR toddlers differ in their proportion of target object fixation? In order to determine if there were any differences between the HR and LR toddlers with respect to the proportion of target fixation, one-way ANOVAs were calculated for the cross-sectional sample that included risk (HR vs. LR) as a between factor for each age point. Results indicated that at 16 months of age the LR group (M = 0.19, SD = 0.06) spent a significantly longer proportion of their fixation duration looking to the target object than the HR group (M = 0.16, SD = 0.06; $F(1, 45) = 4.50$, p = 0.04). In contrast, at 24 months of age, there were no significant differences in the proportion of fixation duration to the target object made by the LR (M = 0.28, SD = 0.09) and HR toddlers (M = 0.26, SD = 0.09; $F(1, 62) = 1.15$, p = .29). This suggests that HR toddlers demonstrate a reduction in preferential looking to the target object compared to LR infants at 16 months of age but catch up by 24 months of age.

To determine whether participants were fixating the target object greater than would be expected by chance, one-sample t-tests were calculated for both the 16 and 24 month groups that compared the proportion of fixation duration of the target object to the value of 1/8 or 0.125. This value of 1/8 assumes equal distribution of fixation duration between all eight items of the array. Since the one-way ANOVA determined that there was a significant group difference at 16 months of age, separate one-sample t-tests were calculated for the LR and HR groups. At 16 months of age both the LR ($t(19) = 4.84$, $p < .001$) and HR ($t(26) = 2.79$, $p = .01$) groups fixated the target object for a proportion of duration greater than chance. This finding indicates that although the LR group fixated the target object proportionally longer than the HR group at 16 months of age, both groups are able to identify and preferentially fixate the target object. Because the one-way ANOVA determined that there were no significant differences in the proportion of fixation duration to the target object between LR and HR toddlers at 24 months of age, the two risk groups were combined. Results revealed that at 24 months of age toddlers spent a significantly greater proportion of time looking to the target object than would be expected by chance ($t(63) = 13.12$, $p < .001$).

Do HR and LR toddlers differ in their proportion of other-category object fixation? Similar to the analyses that were conducted to evaluate preferential target looking, separate one-way ANOVAs examining the proportion of other-category fixation duration were calculated for each age group that included risk (HR vs. LR) as a between subjects factor for the cross-sectional sample. Results indicated that there were no differences by risk group at 16 months of age ($F(1,46) = 1.54$, $p = .22$). Regardless of risk status, toddlers spent approximately 45% ($M = 0.45$, $SD = 0.09$) of their time fixating the four objects that did not belong to the target's superordinate category. A one-sample t-test comparing this mean to the chance value of 4/8 or 0.5 indicated that toddlers were looking to the other-category items significantly less than would be expected assuming equal fixation distribution between all eight objects ($t(46) = -4.36$, $p < .001$).

A group difference trending toward significance was found for other-category object looking at 24 months of age ($F(1, 62) = 3.69$, $p = .06$). At 24 months of age, HR toddlers ($M = 0.40$, $SD = 0.09$) looked longer to the other-category objects than the LR toddlers ($M = 0.36$, $SD = 0.07$). HR toddlers spent 40% of their time looking to the other-category objects while LR toddlers looked to those objects for only 36% of their time. To determine if the participants were fixating the other-category items less than would be expected by chance (0.5) separate one-sample t-test were conducted for LR and HR toddlers at 24 months of age. Results indicated that both LR ($t(41) = -12.53$, $p < .001$) and HR toddlers ($t(21) = -5.70$, $p < .001$) looked to the other-category objects for a duration that was less than chance. This suggests that although there was a trending difference between the LR and HR toddlers, both groups reduced looking to the other-category objects in favor of increasing looking to the target object and/or its other superordinate category members at 24 months of age.

Do HR and LR toddlers demonstrate different developmental trajectories of task performance between 16 and 24 months of age? To test for group differences in developmental trajectories of target fixation, a two-way repeated measures ANOVA was calculated for the proportion of target fixation duration that included age (16-months vs. 24-months) as a within subjects factor and risk (LR vs. HR) as a between subjects factor. A significant main effect for age ($F(1,21) = 31.77$, $p < .001$), indicated that from 16 ($M = 0.17$, $SD = 0.04$) to 24 months of age ($M = 0.27$, $SD = 0.09$), regardless of risk, toddlers increased the proportion of time they spent fixating the target object. There was neither a significant main effect of risk ($F(1,21) = 0.56$, $p < .45$), nor a significant interaction between age and risk ($F(1,21) = 0.18$, $p < .67$).

To test for group differences in developmental trajectories of other-category fixation a two-way repeated measures ANOVA was calculated for the proportion of other-category fixation duration that included age (16-months vs. 24-months) as a within subjects factor and risk (LR vs. HR) as a between subjects factor. A significant main effect for age ($F(1,21) = 16.89$, $p < .001$) indicated that from 16 ($M = 0.47$, $SD = 0.08$) to 24 months of age ($M = 0.38$, $SD = 0.08$), regardless of risk, toddlers decreased the proportion of time they spent fixating the four other-category objects. There was neither a significant main effect of risk ($F(1,21) = 2.70$, $p < .12$), nor a significant interaction between age and risk ($F(1,21) = 0.02$, $p < .90$). These findings suggest that regardless of risk, toddlers display a developmental decline in other-category object fixation between 16 and 24 months of age.

Does performance on this task at 16 months of age correlate with performance at 24 months of age? As was noted for the LR-TD group, the LR group showed a strong association between the proportion of target looking at 16 months of age and the proportion of target looking at 24 months of age ($r(12) = .61$, $p = .02$). No association was found for the HR group ($r(7) = -.14$, $p = .72$).

Do HR and LR toddlers differ in their sequential looking behavior during the Visual Array Task? One-way ANOVAs examining the NRuns target category, the NRuns other category, MRL target category and MRL other category including risk (HR and LR) as a between-subjects factor were calculated for each age to identify any differences in the visual scanning patterns between groups. There were no significant differences identified between the risk groups at either 16 or 24 months of age.

Do patterns of sequential looking change across development? To test for group differences in the longitudinal cohort two-way repeated measures ANOVAs were calculated for NRuns target category, NRuns other-category, MRL target category and MRL other category that included age (16-months vs. 24-months) as a within subjects factor and risk (LR vs. HR) as a between subjects factor. There were no significant main effects or interactions observed for the NRuns other-category or MRL target category variables. A significant main effect of age ($F(1,21) = 9.99$, $p < .01$) as well as a trending main effect of risk ($F(1,21) = 4.07$, $p = .06$) was observed for NRuns target category variable. Regardless of risk, between 16 ($M = 13.65$, $SD = 4.44$) and 24 months of age ($M = 18.74$, $SD = 6.55$), toddlers increased the number of runs that they made within the target category. Regardless of age, HR ($M = 14.06$, $SD = 2.71$) toddlers made significantly fewer runs within the target category than LR toddlers ($M = 17.57$, $SD = 4.73$). There was no interaction observed between risk and age for the NRuns target category variable. A trending main effect of age was also observed for the MRL other category variable ($F(1,21) = 3.39$, $p = .08$). This suggests that, regardless of risk, between 16 ($M = 3.44$, $SD =$

0.90) and 24 (*M* = 2.94, *SD* = 0.55) months of age toddlers may reduce the length of runs between objects that belong to the superordinate category unrelated to the target object. A significant main effect of risk or a risk by age interaction was not observed for the MRL other category variable.

Do the sequential looking behaviors of HR and LR toddlers indicate underlying differences in category knowledge? One-way ANOVAs including risk (HR and LR) as a between-subjects factor were calculated for each age to identify any group differences in the NRun ratio and MRL ratio variables. No differences by risk were found.

Do measures of receptive vocabulary correlate with measures of category knowledge for the HR and LR groups? To assess the relationship between measures of receptive language and object category knowledge during a single testing session, correlation matrices were computed separately for HR and LR toddlers in the cross-sectional sample at each age point. The eye tracking variables included in the matrix were: (1) proportion of fixation duration target, (2) proportion of fixation duration other-category, (3) NRuns ratio, and (4) MRL ratio (see Table 22: LR & Table 23: HR).

At 16 months of age, there were no significant associations between the VAT measure of target label knowledge (proportion of target fixation) and the three measures of category knowledge for the HR infants. There was a significant association between the proportion of target fixation duration and the proportion of other-category fixation duration for the LR infants ($r(18) = -0.46$, $p = .04$). This indicates that as LR toddlers increased looking to the target object, they decreased the amount of time they spent fixating the four other-category objects.

At 24 months of age, there were significant associations between the VAT measure of target label knowledge (proportion of target fixation) and all three measures of category knowledge for both the HR and the LR toddlers. A strong negative association was observed between the proportion of target fixation duration and the proportion of other-category fixation duration for both the LR ($r(40) = -0.79$, $p < .001$) and HR ($r(20) = -0.77$, $p < .001$) toddlers. Again, this indicates a reciprocal relationship between time spent looking to the target object and time spent looking to the four other-category objects. A moderate positive association was observed between the proportion of fixation duration to the target item and the NRuns ratio for the LR ($r(40) = .68$, $p < .001$) and HR ($r(20) = .44$, $p = .04$) toddlers. This finding indicates that as toddlers increase time spent looking to the target item they are also increasing the proportion of runs that they make within the target category. Finally, a moderate to low positive association between the proportion of target fixation duration and the MRL ratio was observed for the LR ($r(40) = .32$, $p = .04$) and HR ($r(20) = .44$, $p = .04$) toddlers. This indicates that as toddlers increase the proportion of time spent looking to the target object they also increase the length of runs they make between objects belonging to the target category compared to items belonging to the other superordinate category.

3.2.1 Summary of risk findings

Both HR and LR toddlers fixated the target object longer than would be expected by chance at 16 and 24 months of age. Early differences in target looking emerged between the risk groups at 16 months of age but dissipated by 24 months of age. This suggests that HR toddlers look to the target item less than their LR peers at 16 months of age but catch up by 24 months of age.

Both HR and LR toddlers fixated the four other-category objects less than would be expected by chance at 16 and 24 months of age. There were no early differences in duration of time looking to the other-category objects. However, there was a trending difference that emerged later. At 24 months of age, even though both groups fixated the other-category items below chance, the HR group fixated those four objects proportionately longer than the LR group. This suggests that the HR toddlers are less likely to inhibit looking to the other-category objects in favor of the target or objects belonging to the same superordinate category as the target. This may be indicative of differences in early category knowledge between the groups. With the exception of the finding that HR toddlers made less runs between target category members (only in the longitudinal cohort), results indicated that there were no differential effects of target label and category knowledge on the sequential looking patterns of HR and LR children.

An early emerging difference in the relationship between target label knowledge and category knowledge for the risk groups was observed at 16 months of age. A reciprocal relationship between target fixation and fixation of the four other-category objects was observed for the LR but not HR group. However, there were no differences between risk groups observed at 24 months of age. For both groups, significant associations were found between the measure of receptive vocabulary and all three measures of target category knowledge. This suggests that while the link between receptive vocabulary and category knowledge may develop earlier in LR groups, HR groups catch up by 24 months of age.

3.3 TODDLERS WITH ATYPICAL DEVELOPMENT

Included in the second aim of this study was the goal to assess any differences in performance on the VAT that may arise between the outcome groups. If you recall, this sample includes data from 44 toddlers collected at 16 months of age ($N_{TD} = 33$, $N_{ASD} = 5$, and $N_{NT} = 6$) and 63 toddlers collected at 24 months of age ($N_{TD} = 51$, $N_{ASD} = 5$, and $N_{NT} = 7$). Embedded in this broader cross-sectional sample is also a longitudinal cohort of 23 toddlers (N_{TD} = 19, N_{ASD} = 1, and N_{NT} = 3). Means and standard deviations for the primary variables of interest may be found in Table 24 (cross-sectional sample) and Table 25 (longitudinal subsample). Due to the small sample sizes of children categorized into the NT and ASD groups only quantitative comparisons of the cross-sectional sample were made. All findings reported below should be considered preliminary and interpreted cautiously.

Do the outcome groups differ in their proportion of target object fixation? To determine if there were any significant differences between toddlers later classified into the TD, NT, or ASD groups regarding their proportion of fixation duration to the target item, one-way ANOVAs were calculated for the cross-sectional sample that included outcome classification (TD vs. NT vs. ASD) as a between subject factor for each age point. Results indicated that at 16 months of age there were no significant differences between groups ($F(2, 41) = 0.13$, $p = .88$). Regardless of classification group inclusion, toddlers spent approximately 17% of their time fixating the target object. This value was greater than would be expected by chance (1/8 or 0.125) ($M = 0.17$, $SD = 0.06$; $t(43) = 5.14$, $p < .001$). Again, this chance value represents the proportion of fixation duration that would be expected if toddlers were to equally distribute their attention among all eight objects within the array.

A group difference was observed at 24 months of age ($F(2, 60) = 3.84$, $p = .03$). Follow-up independent samples t-tests indicated that the ASD group ($M = 0.17$, $SD = 0.04$) looked significantly less to the target object than the TD ($M = 0.28$, $SD = 0.09$; $t(54) = -2.68$, $p = .01$), and NT ($M = 0.30$, $SD = 0.10$; $t(10) = -2.57$, $p = .03$) groups. There was no significant difference in the proportion of target looking measured for the TD and NT groups ($t(56) = -0.65$, $p = .52$). Separate one-sample t-tests were calculated comparing the mean proportion of target looking to the chance value of 0.125 for the ASD group and the combined TD and NT group. While the TD/NT group fixated the target for a longer duration than would be expected by chance ($t(57) = 13.48$, $p < .001$), the ASD group did not ($t(4) = 2.49$, $p = .07$). These findings suggest that group differences in preferentially fixating the target object are not present at 16 months of age but emerge by 24 months of age. At 24 months of age, toddlers later diagnosed with ASD are spending less time looking to the target object than their TD and NT peers as well as no more time than would be expected by chance (equal attention distribution between all eight objects assumed; see Figure 7).

Do the outcome groups differ in their proportion of other-category object fixation? To determine if there were any significant differences between toddlers later classified into the TD, NT, or ASD groups regarding their proportion of fixation duration to the other-category objects, one-way ANOVAs were calculated for the cross-sectional sample that included outcome classification (TD vs. NT vs. ASD) as a between subject factor for each age point. Results indicated that at 16 months of age there were no significant differences between groups ($F(2, 41) = 0.09$, $p = .92$). Regardless of outcome group inclusion, toddlers spent approximately 45% of their time fixating the other-category objects. This value was less than would be expected by chance (4/8; $M = 0.45$, $SD = 0.08$, $t(43) = -4.09$, $p < .001$). Again, this chance value represents the proportion of fixation duration that would be expected if toddlers were to equally distribute their attention among all eight objects within the array.

A group difference was observed at 24 months of age ($F(2, 60) = 4.05$, $p = .02$). Follow-up independent samples t-tests indicated that the ASD group ($M = 0.46$, $SD = 0.10$) looked significantly more to the four other-category objects than the TD ($M = 0.37$, $SD = 0.07$; $t(54) = 2.57$, $p = .01$), and NT groups ($M = 0.34$, $SD = 0.07$; $t(10) = 2.38$, $p = .04$). There was no significant difference in the proportion of target looking measured for the TD and NT groups ($t(56) = -0.97$, $p = .34$). Separate one-sample t-tests were calculated comparing the mean proportion of other-category looking to the chance value of 0.5 for the ASD group and the combined TD and NT group. While the TD/NT group fixated the other-category objects for a shorter duration than would be expected by chance (4/8; $t(57) = -14.52$, $p < .001$), the ASD group did not ($t(4) = -1.01$, $p = .37$). These findings suggest that group differences in fixation of the other-category objects are not present at 16 months of age but emerge by 24 months of age. By 24 months of age, toddlers later diagnosed with ASD are spending more time looking to the other-category objects than their TD and NT peers. Toddlers later diagnosed with ASD were also the only group whose proportion of fixation duration did not differ from the chance value that assumes equal fixation of all objects within the array (see Figure 8).

Do outcome groups differ in their sequential looking behavior during the Visual Array Task? To explore any differences in the tendency to engage in sequential looking between outcome classification groups one-way ANOVAs including outcome (TD, NT, and ASD) as a between-subjects factor were calculated for each age examining NRuns target category, NRuns other-category, MRL target category, and MRL other category. Mirroring the findings for risk comparisons, there were no significant differences identified between the outcome groups at 16 months of age. There was a trending finding identified at 24 months of age signifying a group difference in the MRL target category variable ($F(2, 60) = 2.80$, $p = .07$). Follow-up independent samples t-tests indicated that the ASD group ($M = 2.87$, $SD = 0.53$) made significantly shorter runs within the target category than both the TD ($M = 3.60$, $SD = 0.69$; $t(54) = -2.31$, $p = .03$) and NT groups ($M = 3.52$, $SD = 0.49$; $t(10) = -2.2$, $p = .05$). There was no difference observed between the TD and NT groups ($t(56) = 0.31$, $p = .76$).

To test the differential impacts that category membership may have on scan patterns, one-way ANOVAs for the NRuns ratio and the MRL ratio including outcome (TD, NT, and ASD) as a between-subjects factor were calculated for each age. Once again, no differences by outcome were found at 16 months of age. At 24 months of age a significant difference was identified for the NRuns ratio ($F(2, 60) = 3.58$, $p = .03$) but not the MRL ratio ($F(2, 60) = 1.7$, $p = .19$). Follow-up independent samples t-tests revealed that ASD group ($M = 1.14$, $SD = 0.18$) had a significantly smaller NRuns ratio than both the TD ($M = 1.44$, $SD = 0.42$; $t(54) = -2.95$, $p = .02$) and NT groups ($M = 1.76$, $SD = 0.31$). Additionally, the NT group was trending toward a larger NRuns ratio than the TD group ($t(56) = 1.92$, $p = .06$). This finding suggests that toddlers in the NT group made the greatest proportion of runs within the target category (compared to the other category) followed by those in the TD group and then the ASD group. No difference between groups was reported for the MRL ratio variable.

Do the sequential looking behaviors of outcome groups indicate underlying differences in category knowledge? To determine if the mean values for the NRuns ratio variable reported above reflect a preference for making runs within the target category one-sample t-tests comparing the number of runs ratio to the no difference value of 1 were calculated for each outcome group at each age. It was found that the TD group demonstrated a preference for making runs within the target category at both 16 ($t(31) = 2.31$, $p = .03$) and 24 months of age ($t(50) = 7.51$, $p < .001$). The NT group indicated a preference at 24 ($t(6) = 6.38$, $p < .01$) but not 16 months of age ($t(6) = 0.90$, $p = .40$). And the ASD group did not show a preference for making runs with the target category at either 16 ($t(4) = 1.94$, $p = .12$) or 24 months of age ($t(4) = 1.74$, $p = .16$). Taken together with the findings reported above, this indicates that toddlers in both the TD and NT groups increase the number of runs they make within the target category across age. Toddlers in the TD group demonstrate a preference for making runs between target category members by 16 months of age while toddlers in the NT group demonstrate this preference a bit later at 24 months of age. Toddlers in the ASD group neither show an increase in runs made within the target category nor a preference of sequentially fixating target category members compared to other category members.

Do measures of receptive vocabulary correlate with measures of category knowledge for the HR and LR groups? To assess the relationship between measures of receptive language and object category knowledge during a single testing session, correlation matrices were computed separately for the TD, NT, and ASD toddlers in the cross-sectional sample at each age point. The eye tracking variables included in the matrix were: (1) proportion of fixation duration target, (2) proportion of fixation duration other-category, (3) NRuns ratio, and (4) MRL ratio (see Tables 26-28).

At 16 months of age, similar to the findings from the LR group, a moderate negative association between the proportion of target fixation duration and the proportion of other-category fixation duration was observed for the TD group ($r(31) = -.45$, $p < .01$). Again, this indicated the presence of a reciprocal relationship between target fixation and fixation of the four other-category objects at 16 months of age. There were no significant associations between our measure of target label knowledge (proportion of target fixation) and our three measures of category knowledge for the NT or ASD toddlers.

At 24 months of age, a strong negative association between the proportion of target fixation duration and the proportion of other-category fixation duration was observed for the TD ($r(49) = -.77$, $p < .001$) and NT ($r(5) = -.89$, $p < .01$) groups. No other correlations were observed for the NT group. Additional correlations between the two remaining measures of category knowledge were observed for the TD group. There was a strong positive association between the proportion of target fixation duration and the NRuns ratio ($r(49) = .60$, $p < .001$) as well as a moderate positive association between the proportion of target fixation duration and the MRL ratio ($r(49) = .37$, $p < .01$). This suggests that as toddlers in the TD group increase their looking to the target objects they also increase the proportion of runs they make within the target category as well as the comparative average length of runs within the target category. No significant correlations were identified for the ASD group.

3.3.1 Summary of atypical findings

No differences between outcome groups were observed for the fixation duration variables at 16 months of age. At 24 months of age, the ASD groups demonstrated significantly less target looking and significantly more other-category looking than the TD and NT groups.

Differences in sequential looking between the groups were also observed at 24 but not 16 months of age. First, toddlers in the ASD group were found to make significantly shorter runs between objects of the target category than the toddlers in either the TD or NT group. Second, a number of findings were identified relating to the NRuns ratio. At 24 months of age, there was an ordinal finding that indicated the NT group had the highest NRun ratio followed by the TD and then the ASD group. Both the TD and NT groups displayed an increase in the NRun ratio across age, but the ASD group did not. Finally, values of the NRun ratio for the TD group indicated a preference for sequential looking within the target category at 16 and 24 months of age. This preference was only observed at 24 months of age for the NT group. And was not observed at any age for the ASD group.

Persistent differences in the relationship between receptive vocabulary and category knowledge were observed between outcome groups. While the results for the TD group mirrored those of the LR group, the NT group did not indicate the presence of a relationship between receptive vocabulary and category knowledge until 24 months of age. And even at 24 months of age, only an association between the fixation duration of the target object and the fixation duration of the four other-category objects was noted (an early emerging relationship found for the TD and LR groups). No associations between variables at either 16 or 24 months of age were found for the ASD group.

4.0 DISCUSSION

The primary aims of this research were to (1) validate the visual array task (VAT) as an online measure of receptive vocabulary and category knowledge in a typically developing population and (2) assess group differences in receptive vocabulary and category knowledge using the VAT based on (a) genetic risk for ASD and (b) later outcome classification. Findings pertaining to each of these aims as well as implications for assessment and directions for future research are discussed below.

4.1 TYPICALLY DEVELOPING TODDLERS

Evaluation of the VAT as a measure of receptive vocabulary: *Are toddlers able to accurately identify the target object?* This study was the first to explore the feasibility of expanding the intermodal preferential looking paradigm (IPLP, Golinkoff et al., 1987) to include an eight-item array. As such, a primary concern was whether toddlers would continue to preferentially fixate the target object in the presence of seven distractor items. Despite the added complexity of the task, toddlers as young as 16 months of age were able to accurately identify (preferentially fixate) a known target object embedded within an eight-item array. The successful expansion of this task from the traditional two item pairings of the IPLP and looking-while-listening procedure (LWL, Fernald et al., 1998) to an eight-item array eliminates the concerns surrounding mutual exclusivity and the potential impact of perceptual similarity that are inherent in the forced-choice paradigms. As such, the VAT has the potential to be a more sensitive measure of language comprehension.

Do toddlers demonstrate reduced attention to the other-category items? The expansion of this task to include eight objects also allowed for the simultaneous study of the developmental relationship between receptive vocabulary and category knowledge. In this task it was found that not only did toddlers preferentially fixate the target object, but they also directed their attention away from the other-category members. Both the 16 and 24 month old groups fixated the four other-category objects for a shorter duration than would be expected by chance. This suggests that the target label not only prompted toddlers to fixate the target object, but also lead to less fixation of the other-category objects. This likely indicates an underlying understanding of the superordinate category contrasts tested here, which would be an expected competence for children of these ages (for a review of categorization competencies in infants see Mareschal & Quinn, 2001).

Does the Visual Array Task capture developmental change over time? A developmental improvement in receptive vocabulary is expected between 16 and 24 months of age. If the VAT is able to capture this developmental trajectory, toddlers should demonstrate greater target identification accuracy at 24 months of age compared to 16 months of age. And this is indeed the pattern of results found for the longitudinal cohort. While toddlers fixated the target for a longer duration than would be expected by chance at both 16 and 24 months of age, they fixated the target comparatively longer at 24 months of age. At 16 months of age, toddlers spent approximately 18% of their time fixating the target object. By 24 months of age, this value jumps to 27%.

A developmental improvement in object category knowledge is also expected between 16 and 24 months of age. If the VAT is able to capture this developmental trajectory, toddlers should demonstrate reduced other-category object fixation at 24 months of age compared to 16 months of age. Again, this was the pattern that was identified in the longitudinal cohort. Although toddlers fixated the other-category objects for a shorter duration than would be expected by chance at both 16 and 24 months of age, toddlers spent a significantly shorter amount of time looking to the other-category items at 24 than 16 months of age. At 16 months of age toddlers spent approximately 46% of their time looking to the other-category objects. This dropped to 36% by 24 months of age.

These findings indicate that the VAT was able to capture the expected developmental increases in receptive vocabulary and superordinate category knowledge between 16 and 24 months of age. It also suggests a reciprocal relationship between target looking and other-category looking. As toddlers begin to preferentially fixate the target item they selectively decrease their fixation of the other-category items.

Does performance during the Visual Array Task correlate to two commonly used standardized measures of receptive language? A degree of convergent validity with standardized measures of receptive language has been found for both the IPLP (Golinkoff et al., 2013) and the LWL procedure (Law & Roy, 2008). Thus, it was predicted that the VAT would also demonstrate convergent validity with the two standardized measures collected in this study- the MSEL and the MB-CDI. However, no associations were observed between the measure of target fixation and the subscales of the MSEL or the MB-CDI at either 16 or 24 months of age. Considering the underlying constructs measured by these assessments, this is not all-together surprising.

There are two plausible reasons as to why the measure of target fixation from the VAT did not correlate with the early learning composite (ELC) or the receptive language subscale of the MSEL. First, the MSEL requires that children meet both a basal and a ceiling level of performance for each of the five subscales. This ensures that the assessment captures the full range of developmental capabilities of the child. For the VAT, items were chosen to maximize the likelihood that toddlers would know and therefore preferentially fixate the target objects. It is possible that a number of children reached ceiling level performance on the VAT thus limiting variability. This may have created a discrepancy between the levels of developmental competency measured by the two assessments. Thus, using the procedure with a larger array of test items might yield significant correlations.

The second possibility is that as an experimenter-administered assessment, the receptive language subscale of the MSEL requires a high level of behavioral response from the participant. To demonstrate underlying language comprehension children are asked to respond to prompts such as “where is the door?” or actively point to line drawings in a stimulus book. It is possible that the MSEL is simply under reporting language comprehension abilities compared to the VAT in children of these ages.

With respect to the MB-CDI, there are likely three plausible explanations as to why the measure of target fixation from the VAT did not correlate with the subscales of the MB-CDI. First, only two subscales of the MD-CDI: Words and Gestures (phrases understood and words understood) directly measure language comprehension. The remaining subscales contain elements of expressive language. Second, previous research findings have indicated that parents have a tendency to underreport their children’s language comprehension using the CDI (Houston-Price et al., 2007). And third, again, given the choices made for the first set of stimuli tested with the VAT, it is also possible that a number of children reached ceiling target identification on the VAT.

There were however, negative associations observed between the measure of other-category fixation and the ELC and receptive language subscales of the MSEL at 24 months of age. These associations indicated that an increase in looking to the other-category items was associated with lower overall performance on the MSEL as well as lower receptive language subscales scores. If we conceptualize the measure of other-category looking as an index of both category knowledge and, to some degree, cognitive inhibition (directing attention away from other-category members) then this pattern of findings fits. Toddlers demonstrating higher levels of category knowledge likely also possess higher levels of receptive vocabulary. And those toddlers that are actively inhibiting looking to the other-category items may also demonstrate higher overall levels of cognitive functioning.

Does performance on these variables at 16 months of age correlate with performance at 24 months of age? Correlational data indicated the ability of the VAT to track individual differences in receptive vocabulary between 16 and 24 months of age. Strong correlations between all subscales of the MB-CDI were also noted. No associations were noted for the subscales of the MSEL. This suggests that the VAT may have a predictive value that outperforms the MSEL and is similar to that of the MB-CDI.

Impact of category contrast on attention distribution. This study was also the first to explore whether sequential touching behaviors observed in object manipulation tasks may translate to sequential looking behaviors during a gazed-based task. Results indicated that toddlers do engage in sequential looking behaviors when presented with a visual array of objects that belong to either one of two superordinate categories. There was an overall impact of age indicating that between 16 and 24 months of age, toddlers increased the number of runs made within the target category and decreased the length of runs made between objects of the other-category.

An impact of the verbal target label was also noted. At 16 months of age, children preferentially made sequential visits to objects within the target category. At 24 months of age, toddlers maintained this preference for making runs within the target category but also demonstrated a preference for making shorter runs between other-category objects. Future research should explore the tendency of children to engage in sequential looking in the absence of a verbal prompt. This would more closely mirror traditional tasks of object manipulation where children are presented with eight objects and given no explicit instructions but to engage with the toys (Mandler et al., 1987; Ricciuti, 1965; Starkey, 1981). Although we were able to measure instances of sequential looking, it is possible that hearing the target label lead children to prematurely end sequential looking runs in favor of reorienting toward and fixating the target object. Data reported here may reflect reduced run counts as well as reduced run lengths.

Relationship between receptive vocabulary and category knowledge. Correlation data did not indicate any associations between measures of receptive vocabulary and measures of category knowledge from the VAT at 16 months of age. However, associations between the measure of receptive vocabulary and all three measures of category knowledge were found at 24 months of age. This does not match the findings of Waxman and Markow (1995) that suggest an ability to link nouns and their categories emerges as early as 13 months of age. A lack of findings at 16 months of age may be due to the relative small sample size of LR-TD toddlers (N = 17). A relationship between the proportion of target looking and the proportion of other category looking was observed in the larger LR cohort for the 16-month group.

A 24 months of age, there was a negative association between the proportion of target fixation duration and the proportion of other-category fixation duration. This indicates a reciprocal relationship between the two variables. As toddlers increase their fixation of the target object, they display a corresponding decrease in fixation of the other-category objects. This suggests that by 24 months of age, toddlers have developed a strong enough understanding of the target object and its superordinate category to preferentially fixate the target and its category members. A strong positive association was observed between target fixation and a preference to make runs within the target category as well as the tendency to make longer runs within the target category. Again, this suggests that the target label not only prompts toddlers to look to the target object but also to preferentially direct attention to the other target-category members.

Overall, these findings support the view that receptive vocabulary and category knowledge develop in parallel (Waxman and Markow, 1995). Toddlers that demonstrated greater fixation of the target also demonstrated greater exploration of its corresponding category members through sequential looking behaviors.

4.1.1 Conclusions

The above findings validate the VAT as an online measure of receptive vocabulary for use with typical populations. Although the VAT does not demonstrate convergent validity with the MSEL or the MB-CDI, it does show a strong correlation between performance at 16 and 24 months of age. This consistency in measurement across age points was better than that measured for the MSEL. The VAT was also shown to be a viable online measure of category knowledge. Similar to object manipulation tasks, toddlers demonstrated sequential looking behaviors that were indicative of underlying category knowledge. Moreover, this paradigm allows for an online examination of the emerging relationship between receptive vocabulary and category knowledge through the use of an efficient and simple task.

4.2 TODDLERS AT HIGH- AND LOW-RISK FOR ASD

Accuracy of target identification. An early difference in target fixation emerged between the groups at 16 months of age. The HR toddlers were found to spend proportionately less time fixating the target object than the LR toddlers. However, there was not a significant difference between groups observed at 24 months of age. This suggests that HR infants may be less proficient than their LR peers in identifying the target object at 16 months of age, but catch up by 24 months of age. This may also reflect characteristics of the stimuli used in the VAT. In order to optimize the likelihood of validating the paradigm, the objects included were specifically chosen because they had been rated as commonly understood by toddlers in these age ranges. As task difficulty did not increase between testing sessions, it is possible that the HR group started with a smaller lexicon than the LR group and simply caught up by the second testing session. Given the number of differences in language ability observed between risk groups in the broader literature (Jones et al., 2014; Tager-Flusberg, Paul & Lord, 2005), an underlying difference likely exists between the groups that was not captured by this iteration of the VAT.

No other differences pertaining to target fixation duration were observed between the groups. As predicted, both the HR and LR toddlers fixated the target item for a longer duration than would be expected by chance at 16 and 24 months of age. Furthermore, analyses of the longitudinal cohorts failed to identify differential trajectories of target fixation across age between the HR and LR groups.

These findings appear to be opposite of those previously reported suggesting that a later emerging difference at 36 months of age in target fixation was associated with risk for ASD (Chita-Tegmark at al., 2015). However, the discrepancies in these findings may be due to the differences in age groups tested. In the current study, toddlers were tested with the VAT at 16 and 24 months of age. Chita-Tegmark and colleagues (2015) tested infants at 18, 24, and 36 months of age. Unfortunately, the two ages that report differences between risk groups (16 months of age and 36 months of age) do not overlap between the studies. Future research examining a broader range of developmental time points needs to be conducted in order to reconcile these findings and better delineate a developmental trajectory associated with risk.

Impact of category contrast on attention distribution. Differences between groups for the proportion of other-category fixation followed the opposite pattern. While no differences between risk groups were observed at 16 months of age, a trending difference was observed at 24 months of age. This indicated that toddlers in the HR group fixated the four other-category items longer than toddlers in the LR group. This may be indicative of later emerging differences in category knowledge.

No differences were observed between the groups in the cross-sectional sample with respect to our measures of sequential looking. A main effect was identified in the longitudinal cohort that indicated HR toddlers made fewer runs between target-category members regardless of age. However, they were just as likely as their LR peers to demonstrate a preference for making runs within the target category. This may indicate that the sequential looking behaviors of the HR group are impacted to a lesser degree by hearing the verbal target label. While the HR toddlers still make more runs within the target category than the other-category, this preference does not appear to be as strong.

Relationship between receptive vocabulary and category knowledge. Another early emerging difference between risk groups was observed between the relationship of receptive vocabulary and category knowledge. At 16 months of age, there was a negative association observed between the proportion of target fixation and proportion of other-category fixation for the LR group suggesting a reciprocal relationship between these variables. The emergence of this finding for the LR but not LR-TD group at 16 months of age may simply reflect a lack of power in the LR-TD group due to its limited sample size. No associations between measures of receptive vocabulary and category knowledge were observed for the HR group at 16 months of age.

No differences between risk groups were observed at 24 months of age. Regardless of risk, significant associations between the measure of receptive vocabulary and the three measures of object category knowledge were found for both the LR and HR groups. This finding suggests that the development of the relationship between receptive vocabulary and category knowledge in HR infants may be delayed compared to their LR peers. However, they demonstrate an ability to catch up by 24 months of age, as measured by this iteration of the VAT.

4.2.1 Conclusions

Overall, differences between the risk groups emerged early (16 months of age) and tended to dissipate by 24 months of age. The only exception to this pattern was found for the proportion of other-category fixation. This suggests that the HR group may demonstrate an overall delay in receptive vocabulary and category knowledge but catch up by the second year or that the stimuli chosen for this validation of the VAT do not provide a high enough level of difficulty to induce differences between the two groups at 24 months of age.

4.3 TODDLERS WITH ATYPICAL DEVELOPMENT

Accuracy of target identification. No differences in target looking were observed between outcome groups at 16 months of age. This suggests that early differences in target looking observed for the risk groups may not be indicative of later ASD diagnosis but rather broader autism phenotype characteristics or other unrelated developmental delays such as language delay or global developmental delay. Due to relatively small sample sizes for the ASD and NT groups, these findings should be considered only preliminary.

However, at 24 months of age, the ASD group spent proportionally less time fixating the target object than the TD and NT groups. While the TD and NT groups spent approximately 29% of their time fixating the target object, the ASD group only fixated the target item 17% of the time. This value was no different than chance and did not indicate an increase in target fixation from 16 to 24 months of age. These findings suggest that a lack of developmental increase in looking to the target object may uniquely characterize toddlers that go on to later receive an ASD diagnosis. Again, due to relatively small sample sizes for the ASD and NT groups, these findings should be considered only preliminary.

Impact of category contrast on attention distribution. All differential impacts of category information embedded within the VAT emerged for outcome groups at 24 months of age. The ASD group demonstrated significantly more other-category looking than the TD and NT groups. The TD and NT groups spent only approximately 35% of their time fixating the four other-category objects. The ASD group looked to those objects 46% of the time which was no different than chance. This may either reflect underlying differences in superordinate category knowledge or an inability to inhibit looking to the other-category objects in favor of the target-category objects.

A number of differences in the sequential looking behaviors between groups were also observed. Toddlers with ASD were unique in that they were the only group that failed to demonstrate a preference for making runs within the target category. This preference was established at 16 months for the TD group and 24 months for the NT group. Both the TD and NT groups demonstrated an increase in the preference for making runs within the target category across age. This was not the case for individuals later diagnosed with ASD. The toddlers later diagnosed with ASD also made shorter runs within the target category than toddlers in either the TD or NT group.

Taken together these findings suggest that the scanning patterns of the TD group were most impacted by target label and category knowledge followed by the NT group. There appeared to be no effects of this knowledge on scan patterns for the ASD group. The 24-month group of toddlers later diagnosed with ASD did not appear to perform any better on the task than the 16-month group. This may either reflect a lesser impact of the verbal target label on the scanning patterns of toddlers with ASD or reduced category knowledge.

The VAT was able to capture differences in sequential looking for toddlers later diagnosed with ASD that were not observed in previous studies of sequential touching (Unger & Sigman, 1987; Vitrano, 2015). However, it should be noted that the two previous studies examining this behavior tested older children (3 to 6 years of age). Therefore, it is possible that the difference observed between outcome groups here may dissipate later in development.

Relationship between receptive vocabulary and category knowledge. The findings for the TD group mirrored those of the LR group. At 16 months of age, a negative association was found between the proportion of target fixation and the proportion of other-category fixation, signifying the reciprocal relationship between these two variables. By 24 months of age, the measure of receptive language was correlated with all three measures of category knowledge.

There were no associations between measures observed for the NT group at 16 months of age. However, by 24 months of age a negative association between the proportion of target fixation and the proportion of other-category fixation emerged. This suggests that a relationship between object label knowledge and category knowledge may be delayed in toddlers of the NT group.

No associations between variables at either 16 or 24 months of age were found for the ASD group. This again suggests, at the least, a delay in the development of a relationship between object label knowledge and category knowledge for toddlers in the ASD group. These findings mirror those of Vitrano (2015) and not Ungerer and Sigman (1987).

4.3.1 Conclusions

Overall, toddlers later diagnosed with ASD displayed a significant reduction in performance on the VAT compared to the TD and NT groups. At 24 months of age they demonstrated a reduced accuracy in identifying the target object and the category contrasts embedded in the task had little influence on their attention distribution. There were also no measured associations between receptive vocabulary and category knowledge. But perhaps the most defining feature of the ASD group was a lack of an increase in ability observed between the 16 and 24 month old groups.

4.4 GENERAL DISCUSSION

4.4.1 The VAT as an assessment measure

The development of behavioral measures of receptive language such as the intermodal preferential looking paradigm (IPLP; Golinkoff et al., 1987) and the looking-while-listening procedure (LWL; Fernald et al., 1998) have provided researchers with an objective measure of language comprehension. These tasks require a lesser behavioral response from participants compared to more commonly utilized developmental assessments such as the Bayley Scales of Infant Development (Bayley, 1993) and the Mullen Scales of Early Learning (MSEL; Mullen, 1995). This reduction in task demands makes procedures like the IPLP and the LWL ideal for use with young or atypically developing populations. It has been suggested that these methods may also provide a more accurate measure of receptive vocabulary than parent report measures (Houston-Price et al., 2007). Despite reducing the impact of confounds associated with interaction-based assessments and parent-report questionnaires, these attention-based measures of receptive language still have limitations. Chiefly, both the IPLP and the LWL procedure, with few exceptions (Brady et al., 2014), rely on a forced-choice paradigm. When a target item is paired with only a single distractor the interpretation of task performance is heavily reliant upon the chosen stimuli. For example, if a target item (e.g. hat) is paired with an item belonging to another functional category (e.g. banana) the participant might not identify the target item based on vocabulary knowledge but rather broader category knowledge. If care is not taken to select appropriate stimulus sets, this may lead to an overestimation of participant ability.

The present study was the first to explore the feasibility of expanding the IPLP to an eight-item array, referred to here as the visual array task (VAT). As predicted, toddlers were able to successfully identify a labeled target object from among seven distractors at both 16 and 24 months of age and demonstrated a developmental increase in target identification across age. Furthermore, a strong correlation was found between performance at 16 months of age and 24 months of age suggesting consistency of measurement over time. In fact, this association was stronger for the measure of target identification derived from the VAT than the receptive language subscale of the MSEL. This confirms the feasibility of using visual attention to measure receptive vocabulary outside of the forced-choice paradigm. Successfully identifying the target object (e.g. hat) from both same category members (e.g. pants, a shirt, and a shoe) and other-category members (e.g. a banana, bread, an apple or a cookie) requires a more accurate understanding of word-comprehension than identifying a target object from a single alternative. Therefore, the VAT has the potential to provide a more confident assessment of receptive vocabulary by demonstrating that toddlers are able to reliably identify target objects from a larger array of items containing both category and non-category members.

Evidence from visual paired preference tasks suggest that a sensitivity to superordinate categories develops around three to four months of age (Mareschal & Quinn, 2001). While the current study did not test infants that young, data does indicate an understanding of superordinate category membership at 16 months of age. An increase in object category knowledge was also captured between 16 and 24 months of age. As discussed above, as toddlers spent more time looking to the target object, they reduced the amount of time spent looking to the other-category objects. This suggests that labeling the target object draws visual attention to the target as well as categorically similar objects within the array, thus indicating a conceptual understanding of their category relationship.

By expanding the visual stimuli to an eight item array, the VAT also allowed for the examination of category understanding through a measure of sequential looking. Findings from Mandler and Bauer (1988) suggest that toddlers as young as 20 months of age will sequentially touch members of the same category when presented with eight miniatures representing objects belonging to two distinct superordinate categories (e.g. animals versus vehicles). The current study was the first to explore whether this tendency to engage in sequential touching of category members would also extend to patterns of sequential looking between category members. Unfortunately, unlike traditional tasks of object manipulation, this task did not provide participants with the opportunity to freely explore (visually scan) the objects on the screen prior to target label onset. Thus, the scan patterns of participants were likely influenced by knowledge of the target category. In a way, this transformed the task from the “free-play” encouraged in object manipulation paradigms to a form of visual search. However, an indication of superordinate category knowledge was reflected in a preference for sequentially looking to two or more objects belonging to the target category at 16 months of age. By 24 months of age an underlying understanding of superordinate categories was also indicated by a preference for making longer sequences of looks between target category members than other-category members. This suggests that despite potential interference from the target label, toddlers do engage in sequential looking behaviors akin to the sequential touching behaviors observed in object manipulation tasks.

The ability of the VAT to measure both receptive vocabulary and object category knowledge provides a unique opportunity to examine the developing relationship between these two competencies within the context of the same task. Fulkerson and Waxman (2007) assert that as toddlers expand their receptive vocabularies they are also better able to identify commonalities between objects belonging to the same category. Data reported here support this argument. The reciprocal relationship between the distribution of attention to the target object and the other-category items suggests that as word-comprehension increases so does category knowledge. Significant correlations indicating a relationship between measures of target identification and category knowledge observed at 24 months of age further bolster this claim. Overall, results from the typically developing sample indicated that the tested visual array task has the potential to measure receptive vocabulary as well as category knowledge within the context of a single gaze-based assessment.

4.4.2 Possible applications of the VAT in atypically developing populations

Previous research has reported early delays in receptive language for children that later go on to receive an ASD diagnosis (Zwaigenbaum et al., 2005; Landa & Garrett-Mayer, 2006; Mitchell et al. 2006). A difference in target identification between the risk groups was observed at 16 but not 24 months of age. One interpretation of this findings may be that toddlers at high-risk (HR) for ASD demonstrate early delays in receptive vocabulary but quickly catch up to their low-risk (LR) peers. However, it is more likely the case that the current iteration of the VAT did not include more advanced vocabulary words that would have differentiated the two groups. This limitation as well as how it may be addressed in the future is further discussed in the following section.

This study was the first to examine the development of object category knowledge in a HR and LR infant sibling population. Research with older children and adults with ASD has suggested a general deficit in the domain of categorization (Gastgeb & Strauss, 2012) as well as a weakened link between receptive language abilities and category knowledge (Ungerer &

Sigman1987). While a trending finding suggested a lesser influence of the target label on attention distribution for the HR group compared to the LR group, this did not reach a level of significance. Therefore, data reported here do not support a deficit in the early knowledge of superordinate categories.

A significant pattern observed was a weakened relationship between the measure of receptive language (target identification) and measures of superordinate category understanding (fixation of other-category members and patterns of sequential looking) at 16 months of age. This suggests that differences in category knowledge may be a meaningful measure to consider when conceptualizing the receptive language deficits associated with HR populations. During language learning, typically developing children are sensitive to the common features of category members and use this information to generalize known object labels to newly encountered exemplars (Ferguson & Waxman, 2017). The inability to identify similarities between category members and/or utilize this information when learning new object referents may be a potential mechanism for the delays observed in the HR population. A lack of a similar finding at 24 months of age may simply be due to the limited vocabulary tested during this particular session. Future studies should further explore this emerging relationship in children at risk for and with ASD.

Although this study had a limited number of toddlers that went on to receive a formal diagnosis of ASD, data indicated that a lack of developmental increase in task performance may be uniquely indicative of the disorder. While both the typically developing (TD) and atypically developing (non-ASD) groups demonstrated an increase in receptive vocabulary (target identification) and category knowledge over time, the ASD group did not. This may represent either a more severe delay or a plateau of early receptive vocabulary development for these toddlers. Future longitudinal studies should include both a larger population of children later diagnosed with ASD as well as more frequent points of measurement across a wider developmental window to distinguish the broader developmental trajectory.

Similar to the findings reported for the HR group, no association between receptive vocabulary and categorization abilities for the toddlers later diagnosed with ASD was noted. However, unlike the HR group, a lack of this relationship was observed at both 16 and 24 months of age. Again, this finding points to an inability to generalize target labels from specific exemplars to newly encountered category members as a possible underlying mechanism for the delays in receptive vocabulary observed in children with ASD. It should also be noted that children categorized as having ASD here likely reflect a subset of the broader ASD population with more severe symptomology. The average age of diagnosis in a community sample of children with ASD (where data was aggregated at 8 years of age) is 52 months (Baio et al., 2018). In order to receive a diagnosis at 48, 36 or even 24 months of age, children must display clearly discernable levels of ASD symptomology. Therefore, this may be indicative of only children more severely affected with ASD or those with ASD and associated language impairment. Again, future studies should include data from a larger and more representative population.

The low behavioral demands placed on participants makes the VAT an ideal measure to use with atypically developing or otherwise impaired populations where task demands may be prohibitory to the successful completion of current standardized measures. Data reported here validate the feasibly of using the VAT to measure receptive vocabulary and category knowledge in children at genetic risk for ASD. Without modification, the VAT can accommodate the testing of non-verbal to minimally verbal populations, as well as populations, such as children with anxiety, that may encounter many of the interaction-based hurdles similar to children with ASD. Thus, the VAT is measure that can be administered uniformly across a broad spectrum of populations.

4.4.3 Limitations of the present study and future directions

This research is an exciting first step toward developing a more sensitive attention based measure of receptive vocabulary. As such, the data produced by this preliminary version of the task has a number of limitations. The present study was designed as a test of procedural concept. In order to determine whether or not children could successfully identify a target object from a larger array of eight items, the objects included in the stimulus were carefully chosen to represent vocabulary words that children of the tested age ranges would be reasonably expected to know. While this increased the likelihood that toddlers would indeed preferentially fixate the target object within the larger eight item array, it also may have resulted in a number of participants reaching ceiling performance. This reduction in overall variability may have lead to the inability of the task to capture differences in target identification at 24 months of age as well as any differences in the developmental trajectories of object label and category knowledge between risk groups. Future iterations of the task should include both known and unknown vocabulary words.

In order to better mimic the object manipulation paradigm, future testing of the VAT should also include a delay between the onset of the visual stimuli and the labeling of the target object. This will allow for a baseline measure of visual scan patterns without interference from a target label. Due to the simultaneous presentation of the object array and target label, it is not yet possible to infer exactly how the sequential scanning of participants may have been influenced by differential knowledge of the target word.

Additionally, the task was comprised of only 12 stimulus trials. Despite the limited number of words and category contrasts tested, a number of interesting results were produced that validate the potential of the task. The VAT could easily be expanded into a more exhaustive assessment measure. By shortening the presentation length of each trial to five seconds, upwards of 100 words could theoretically be tested during a ten-minute session. This represents a significant reduction in time required for testing compared to existing measures of receptive language such as the MSEL which can take upwards of an hour to administer. By manually advancing from one trial to the next, experimenters also have the flexibility to present trials at a pace that is tailored to each participant and take breaks in the testing procedure as needed.

As is true of the IPLP and the LWL procedure, the VAT also allows for flexibility in the design of its stimulus arrays. The particular iteration of the VAT used in this study tested receptive vocabulary knowledge, or to be more precise, object noun labels, as well as a superordinate category contrast. However, the general design of the task allows for the flexibility to test any number of lexical constructs or category contrasts.

Despite its limitations, results indicate that the tested visual array task may be both a sensitive measure of receptive vocabulary as well as capable of reflecting gains in category knowledge. This paradigm has the potential to be developed into an inexpensive and efficient task of receptive vocabulary as well as other general constructs such as category knowledge that places few behavioral demands on the participant. Once this task is developed into a standardized form it may be particularly useful in testing for receptive vocabulary or category understanding delays in atypically developing or developmentally delayed populations. As it requires no overt gestural or expressive language response, it can be administered across the developmental spectrum including preverbal infants and non-verbal adults.

5.0 TABLES

Table 1. Demographic information for the longitudinal cohort split by risk

	Low-Risk (N = 14)		High-Risk (N = 9)	
Chronological Age (months)				
16-month time point (*M*, *SD*)	14	17.04, 1.29	9	17.09, 1.10
24-month time point (*M*, *SD*)	14	24.79, 0.60		25.82, 1.27
Gender				
Male (%)	10	(71%)	7	(78%)
Female (%)	4	(29%)	2	(22%)
Racial or Ethnic Minority (%)	1	(7%)	1	(11%)
Maternal Education				
High School (%)	0	(0%)	1	(25%)
Some College or College Degree (%)	4	(36%)	1	(25%)
Graduate of Professional School (%)	7	(64%)	2	(50%)
Paternal Education				
High School (%)	0	(0%)	0	(0%)
Some College or College Degree (%)	6	(43%)	5	(56%)
Graduate of Professional School (%)	8	(57%)	4	(44%)

Table 2. Standardized assessment scores for the longitudinal cohort split by risk

	Low-Risk			High-Risk		
	N	*M*	*SD*	*N*	*M*	*SD*
16 Months						
ADOS						
Severity Score	14	1.36	0.84	9	2.22	2.11
Communication	14	0.50	0.65	9	0.89	1.62
Social	14	0.57	0.76	9	2.89	2.62
Behavior	14	0.93	1.07	9	0.67	1.00
Mullen Scales of Early Learning						
Early Learning Composite	13	99.54	12.16	8	96.50	17.56
Verbal Composite	13	100.77	16.96	9	93.33	23.52
Non-Verbal Composite	13	108.23	9.98	8	106.75	10.94
Receptive Language						
Raw Score	13	17.77	2.80	9	16.33	3.50
T-Score	13	48.31	11.18	9	43.78	13.08
Expressive Language						
Raw Score	13	15.54	2.63	9	15.11	4.29
T-Score	13	47.15	9.50	9	46.33	14.57
MB-CDI Words and Gestures						
Phrase						
Number	14	20.64	7.98	9	17.78	7.97
Percentile	14	52.57	29.16	9	35.67	29.42
Understands						
Number	14	173.50	104.81	9	146.78	104.02
Percentile	14	43.21	33.52	9	33.56	33.07
Understands and Says						
Number	14	34.64	27.74	9	32.89	41.44
Percentile	14	42.21	24.56	9	31.56	27.09

24 Months						
ADOS						
Severity Score	14	1.36	0.84	9	2.22	2.11
Communication	14	0.50	0.65	9	0.89	1.62
Social	14	0.57	0.76	9	2.89	2.62
Behavior	14	0.93	1.07	9	0.67	1.00
Mullen Scales of Early Learning						
Early Learning Composite	13	110.69	13.97	8	97.63	17.07
Verbal Composite	13	114.23	14.00	9	103.56	22.26
Non-Verbal Composite	13	109.92	18.12	8	98.50	11.15
Receptive Language						
Raw Score	14	27.36	2.56	9	25.22	4.52
T-Score	14	59.07	7.46	9	52.33	13.92
Expressive Language						
Raw Score	14	24.21	3.75	9	22.67	4.44
T-Score	14	53.57	9.76	9	49.67	12.59
MB-CDI Words and Sentences						
Words Produced						
Number	11	358.82	198.73	9	246.33	191.37
Percentile	11	58.73	31.43	9	40.44	30.56
Length of Three Longest Sentences						
Average	11	4.03	1.69	9	3.18	1.95
Percentile	11	51.09	25.15	9	35.67	28.26
Complexity						
Number	11	11.18	10.21	9	6.11	10.08
Percentile	11	67.64	19.44	9	46.67	27.31

Table 3. Demographic information for the cross-sectional sample split by risk

	Low-Risk (N = 62)		High-Risk (N = 49)	
16 Months				
Chronological Age in months (M, SD)	20	16.97, 1.9	27	16.73, 0.77
Gender				
Male (%)	13	(65%)	17	(63%)
Female (%)	7	(35%)	10	(37%)
Racial or Ethnic Minority (%)	1	(5%)	7	(26%)
Maternal Education				
High School (%)	0	(0%)	3	(20%)
Some College or College Degree (%)	6	(43%)	6	(40%)
Graduate of Professional School (%)	8	(57%)	6	(40%)
Paternal Education				
High School (%)	0	(0%)	2	(8%)
Some College or College Degree (%)	7	(39%)	13	(52%)
Graduate of Professional School (%)	11	(61%)	10	(40%)
24 Months				
Chronological Age in months (M, SD)	42	24.75, 0.71	22	25.46, 1.21
Gender				
Male (%)	25	(60%)	13	(59%)
Female (%)	17	(40%)	9	(51%)
Racial or Ethnic Minority (%)	4	(10%)	3	(14%)
Maternal Education				
High School (%)	0	(0%)	1	(9%)
Some College or College Degree (%)	12	(46%)	6	(55%)
Graduate of Professional School (%)	14	(54%)	4	(36%)
Paternal Education				
High School (%)	0	(0%)	0	(0%)
Some College or College Degree (%)	24	(60%)	12	(60%)
Graduate of Professional School (%)	16	(40%)	8	(40%)

Table 4. Standardized assessment scores for the cross-sectional sample split by risk

	Low-Risk			High-Risk		
	N	*M*	*SD*	*N*	*M*	*SD*
16 Months						
Chronological Age	20	16.97	1.09	27	16.73	0.77
ADOS						
Severity Score	19	1.47	1.12	25	2.36	1.89
Communication	19	0.68	1.06	25	1.60	1.98
Social	19	0.79	1.23	25	2.64	2.69
Behavior	19	0.79	0.98	25	0.92	0.95
Mullen Scales of Early Learning						
Early Learning Composite	19	98.63	10.49	25	94.40	14.31
Verbal Composite	19	100.00	14.89	25	92.44	20.84
Non-Verbal Composite	19	107.05	8.90	24	107.50	9.00
Receptive Language						
Raw Score	19	17.92	2.99	26	16.15	3.88
T-Score	19	48.47	10.99	26	42.69	13.90
Expressive Language						
Raw Score	19	15.47	2.44	26	14.42	3.42
T-Score	19	46.79	9.40	26	43.42	12.08
MB-CDI Words and Gestures						
Phrase						
Number	20	20.55	7.56	24	19.25	6.68
Percentile	20	51.00	29.54	24	40.83	29.01
Understands						
Number	20	161.90	97.18	24	150.13	86.07
Percentile	20	38.90	29.38	24	34.29	28.37
Understands and Says						
Number	20	34.60	32.15	24	28.63	33.44
Percentile	20	39.15	26.12	24	30.25	24.47

24 Months						
Chronological Age	42	24.75	0.71	22	25.46	1.21
ADOS						
Severity Score	40	1.28	0.64	22	2.50	1.92
Communication	40	0.88	0.88	22	1.23	1.31
Social	40	1.18	1.55	22	2.59	2.30
Behavior	40	0.60	0.90	22	1.14	1.39
Mullen Scales of Early Learning						
Early Learning Composite	40	112.27	16.02	18	90.94	16.70
Verbal Composite	39	113.97	20.85	19	93.32	25.95
Non-Verbal Composite	39	111.79	17.74	20	94.95	9.57
Receptive Language						
Raw Score	41	26.98	3.09	20	22.75	6.55
T-Score	41	58.68	9.29	20	46.55	16.51
Expressive Language						
Raw Score	41	24.34	4.86	20	20.60	5.36
T-Score	41	54.27	11.55	20	44.25	14.21
MB-CDI Words and Sentences						
Words Produced						
Number	35	313.03	165.43	20	212.10	182.80
Percentile	35	52.29	27.42	20	34.70	29.61
Length of Three Longest Sentences						
Average	35	3.99	1.72	20	3.12	2.12
Percentile	35	51.26	25.31	20	33.60	28.37
Complexity						
Number	35	8.06	8.60	20	4.85	9.24
Percentile	35	62.06	21.79	20	44.00	22.03

Table 5. Demographic information for the longitudinal cohort split by outcome

	TD (*N* = 19)		ASD (*N* = 1)		NT (*N* = 3)	
Chronological Age (months)						
16-month time point (*M, SD*)	19	16.91, 1.16	1	18.50, -	3	17.53, 1.37
24-month time point (*M, SD*)	19	25.12, 0.96	1	27.40, -	3	24.93, 0.93
Gender						
Male (%)	15	(79%)	1	(100%)	1	(33%)
Female (%)	4	(21%)	0	(0%)	2	(67%)
Racial or Ethnic Minority (%)	2	(11%)	0	(0%)	0	(0%)
Maternal Education						
High School (%)	1	(8%)	0	(0%)	0	(0%)
Some College or College Degree (%)	4	(33%)	1	(100%)	0	(0%)
Graduate of Professional School (%)	7	(58%)	0	(0%)	2	(100%)
Paternal Education						
High School (%)	0	(0%)	0	(0%)	0	(0%)
Some College or College Degree (%)	9	(47%)	0	(0%)	2	(67%)
Graduate of Professional School (%)	10	(53%)	1	(100%)	1	(33%)

Table 6. Standardized assessment scores for the longitudinal cohort split by outcome

	TD			ASD			NT		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
16 Months									
Chronological Age	19	16.91	1.16	1	18.50	-	3	17.53	1.37
ADOS									
Severity Score	19	1.26	0.73	1	7.00	-	3	2.67	1.53
Communication	19	0.47	0.61	1	5.00	-	3	0.33	0.58
Social	19	0.74	0.87	1	8.00	-	3	4.00	1.73
Behavior	19	0.68	0.95	1	3.00	-	3	1.00	1.00
Mullen Scales of Early Learning									
Early Learning Composite	17	100.82	13.04	1	70.00	-	3	94.00	12.12
Verbal Composite	18	101.11	18.02	1	59.00	-	3	90.33	19.76
Non-Verbal Composite	17	109.35	9.44	1	91.00	-	3	103.67	11.02
Receptive Language									
Raw Score	18	17.61	2.89	1	14.00	-	3	15.67	4.51
T-Score	18	48.50	11.32	1	34.00	-	3	38.33	13.58
Expressive Language									
Raw Score	18	15.78	2.86	1	7.00	-	3	15.67	3.06
T-Score	18	48.39	10.60	1	20.00	-	3	46.33	7.51
MB-CDI Words and Gestures									
Phrase									
Number	19	20.47	7.45	1	9.00	-	3	17.00	10.82
Percentile	19	51.05	28.55	1	1.00	-	3	28.67	27.02
Understands									
Number	19	173.68	101.92	1	39.00	-	3	137.00	115.58
Percentile	19	44.32	33.35	1	1.00	-	3	21.33	21.60
Understands and Says									
Number	19	36.37	34.36	1	9.00	-	3	27.00	28.58
Percentile	19	42.53	25.82	1	9.00	-	3	19.33	5.508

24 Months									
Chronological Age	19	25.12	0.96	1	27.40	-	3	24.93	0.93
ADOS									
Severity Score	19	1.26	0.73	1	7.00	-	3	2.67	1.53
Communication	19	0.47	0.61	1	5.00	-	3	0.33	0.58
Social	19	0.74	0.87	1	8.00	-	3	4.00	1.73
Behavior	19	0.68	0.95	1	3.00	-	3	1.00	1.00
Mullen Scales of Early Learning									
Early Learning Composite	17	107.71	13.19	1	81.00	-	3	102.67	29.50
Verbal Composite	18	113.22	13.41	1	79.00	-	3	100.00	34.77
Non-Verbal Composite	17	106.65	17.03	1	85.00	-	3	106.33	13.58
Receptive Language									
Raw Score	19	27.05	2.68	1	25.00	-	3	23.67	7.51
T-Score	19	58.11	7.39	1	48.00	-	3	48.67	24.83
Expressive Language									
Raw Score	19	24.21	3.49	1	16.00	-	3	22.33	5.69
T-Score	19	53.68	8.96	1	30.00	-	3	49.00	17.10
MB-CDI Words and Sentences									
Words Produced									
Number	16	341.94	191.86	1	13.00	-	3	226.67	191.17
Percentile	16	56.88	29.94	1	1.00	-	3	33.00	27.73
Length of Three Longest Sentences									
Average	16	4.04	1.75	1	1.00	-	3	2.43	1.25
Percentile	16	51.06	24.63	1	1.00	-	3	21.67	18.01
Complexity									
Number	16	9.31	9.82	1	0.00	-	3	9.67	15.04
Percentile	16	61.81	22.82	1	20.00	-	3	51.67	34.03

Table 7. Demographic information for the cross-sectional sample split by outcome

	TD		ASD		NT	
	(*N* = 84)		(*N* = 10)		(*N* = 13)	
16 Months						
Chronological Age (months)	33	16.74, 0.93	5	17.06, 1.01	6	17.15, 1.04
Gender						
Male (%)	21	(64%)	5	(100%)	3	(50%)
Female (%)	12	(36%)	0	(0%)	3	(50%)
Racial or Ethnic Minority (%)	4	(12%)	2	(40%)	2	(33%)
Maternal Education						
High School (%)	1	(5%)	1	(25%)	1	(33%)
Some College or College Degree (%)	10	(48%)	1	(25%)	0	(0%)
Graduate of Professional School (%)	10	(48%)	2	(50%)	2	(67%)
Paternal Education						
High School (%)	1	(3%)	1	(33%)	0	(0%)
Some College or College Degree (%)	13	(41%)	1	(33%)	4	(67%)
Graduate of Professional School (%)	18	(56%)	1	(33%)	2	(33%)
24 Months						
Chronological Age (months)	51	24.97, 0.92	5	25.14, 1.34	7	25.16, 0.82
Gender						
Male (%)	30	(59%)	4	(80%)	3	(43%)
Female (%)	21	(41%)	1	(20%)	4	(57%)
Racial or Ethnic Minority (%)	3	(6%)	2	(40%)	2	(29%)
Maternal Education						
High School (%)	1	(3%)	0	(0%)	0	(0%)
Some College or College Degree (%)	15	(48%)	2	(67%)	1	(33%)
Graduate of Professional School (%)	15	(48%)	1	(33%)	2	(67%)
Paternal Education						
High School (%)	0	(0%)	0	(0%)	0	(0%)
Some College or College Degree (%)	30	(61%)	1	(25%)	5	(71%)
Graduate of Professional School (%)	19	(39%)	3	(75%)	2	(29%)

Table 8. Standardized assessment scores for the cross-sectional sample split by outcome

	TD			ASD			NT		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
16 Months									
Chronological Age	33	16.74	0.93	5	17.06	1.01	6	17.15	1.04
ADOS									
Severity Score	33	1.24	0.61	5	5.80	0.84	6	2.83	1.17
Communication	33	0.55	0.75	5	5.00	1.00	6	1.67	1.63
Social	33	0.85	1.20	5	6.80	1.64	6	3.17	1.72
Behavior	33	0.64	0.86	5	1.80	0.84	6	1.33	1.03
Mullen Scales of Early Learning									
Early Learning Composite	30	99.87	11.16	5	81.40	10.21	6	96.00	8.85
Verbal Composite	31	99.48	16.68	5	69.60	17.10	6	95.00	17.32
Non-Verbal Composite	30	108.33	8.95	5	105.00	10.12	6	104.00	9.25
Receptive Language									
Raw Score	31	17.42	3.03	5	13.00	2.65	6	16.33	4.03
T-Score	31	47.84	11.45	5	31.00	10.22	6	41.83	12.42
Expressive Language									
Raw Score	31	15.42	2.78	5	11.40	3.21	6	16.00	2.37
T-Score	31	47.26	10.30	5	32.80	11.43	6	47.83	6.18
MB-CDI Words and Gestures									
Phrase									
Number	33	20.61	6.55	3	9.00	1.00	6	18.50	7.15
Percentile	33	48.97	28.22	3	6.33	4.73	6	31.83	18.28
Understands									
Number	33	167.67	90.98	3	34.67	8.39	6	137.33	76.17
Percentile	33	41.79	29.52	3	1.00	0.00	6	24.67	17.06
Understands and Says									
Number	33	34.06	34.29	3	6.67	2.08	6	31.17	34.49
Percentile	33	38.64	38.64	3	12.67	3.22	6	26.83	24.00

24 Months									
Chronological Age	51	24.97	0.92	5	25.14	1.34	7	25.16	0.82
ADOS									
Severity Score	51	1.28	0.66	5	5.40	1.14	6	2.33	1.21
Communication	51	0.80	0.80	5	2.60	1.52	6	1.33	1.50
Social	51	1.02	1.14	5	5.40	1.52	6	4.17	2.48
Behavior	51	0.59	0.85	5	3.20	1.10	6	0.50	0.84
Mullen Scales of Early Learning									
Early Learning Composite	47	109.06	15.54	4	71.50	7.77	6	101.50	27.50
Verbal Composite	48	112.90	18.15	4	56.00	16.83	6	95.83	31.47
Non-Verbal Composite	47	107.34	16.90	5	89.80	5.97	7	109.29	21.24
Receptive Language									
Raw Score	49	26.67	3.19	4	14.75	7.59	7	24.00	6.22
T-Score	49	57.55	9.53	4	28.00	13.47	7	49.43	19.36
Expressive Language									
Raw Score	49	24.37	4.30	5	13.20	1.92	6	21.00	6.00
T-Score	49	54.24	10.05	5	25.00	4.12	6	45.50	17.43
MB-CDI Words and Sentences									
Words Produced									
Number	44	306.11	171.74	4	28.25	33.82	7	230.86	143.55
Percentile	44	51.55	27.96	4	3.75	5.50	7	34.43	20.57
Length of Three Longest Sentences									
Average	44	4.07	1.82	4	1.00	0.00	7	2.69	1.36
Percentile	44	51.80	24.67	4	1.00	0.00	7	26.14	20.81
Complexity									
Number	44	7.66	9.03	4	0.00	0.00	7	6.00	9.49
Percentile	44	58.50	23.70	4	32.50	8.66	7	49.71	21.07

Table 9. Objects belonging to each superordinate category

Superordinate Category	*Object Category Members*			
Vehicles	Train	Airplane	Car	Truck
Clothing	Hat	Pants	Shoe	Shirt
Animals	Cat	Dog	Horse	Bird
Food	Bread	Banana	Cookie	Apple
Furniture	Door	Table	Bed	Chair
Utensils	Plate	Cup	Spoon	Bottle

Table 10. Demographic information for the low-risk, typically developing cross-sectional sample

	16 Months (N = 17)		24 Months (N = 37)	
Chronological Age in months (*M*, *SD*)	17	16.94, 1.18	37	24.74, 0.73
Gender				
Male (%)	12	(71%)	22	(59%)
Female (%)	5	(29%)	15	(41%)
Racial or Ethnic Minority (%)	1	(6%)	2	(5%)
Maternal Education				
High School (%)	0	(0%)	0	(0%)
Some College or College Degree (%)	6	(50%)	11	(46%)
Graduate of Professional School (%)	6	(50%)	13	(54%)
Paternal Education				
High School (%)	0	(0%)	0	(0%)
Some College or College Degree (%)	5	(31%)	21	(58%)
Graduate of Professional School (%)	11	(69%)	15	(42%)

Table 11. Standard assessment scores for the low-risk, typically developing cross-sectional sample

	16 Months			24 Months		
	N	*M*	*SD*	*N*	*M*	*SD*
Chronological Age	17	16.94	1.18	37	24.75	.73
Mullen Scales of Early Learning						
Early Learning Composite	16	99.56	11.03	35	113.37	14.82
Verbal Composite	16	100.38	16.10	35	115.80	19.16
Non-Verbal Composite	16	107.81	8.77	35	111.66	16.94
Receptive Language						
Raw Score	16	17.56	2.80	36	27.25	2.57
T-Score	16	47.69	10.95	36	59.53	7.73
Expressive Language						
Raw Score	16	15.81	2.46	36	24.75	4.68
T-Score	16	48.13	9.19	36	55.31	10.66
MB-CDI Words and Gestures						
Phrase						
Number	17	20.82	7.63	-	-	-
Percentile	17	53.24	29.36	-	-	-
Understands						
Number	17	168.35	97.56	-	-	-
Percentile	17	41.82	29.96	-	-	-
Understands and Says						
Number	17	37.82	33.80	-	-	-
Percentile	17	43.06	26.44	-	-	-
MB-CDI Words and Sentences						
Words Produced						
Number	-	-	-	31	323.32	168.51
Percentile	-	-	-	31	54.48	27.87
Length of Three Longest Sentences						
Average	-	-	-	31	4.15	1.70
Percentile	-	-	-	31	54.19	24.26
Complexity						
Number	-	-	-	31	7.90	8.30
Percentile	-	-	-	31	62.39	22.03

Table 12. Demographic information for the low-risk, typically developing longitudinal cohort

	LR-TD (*N* = 13)	
Chronological Age 16 months (*M*, *SD*)	13	17.05, 1.34
Chronological Age 24 months (*M*, *SD*)	13	24.82, 0.61
Gender		
Male (%)	10	(77%)
Female (%)	3	(23%)
Racial or Ethnic Minority (%)	1	(8%)
Maternal Education		
High School (%)	0	(0%)
Some College or College Degree (%)	4	(40%)
Graduate of Professional School (%)	6	(60%)
Paternal Education		
High School (%)	0	(0%)
Some College or College Degree (%)	5	(38%)
Graduate of Professional School (%)	8	(62%)

Table 13. Standardized assessment scores for the low-risk, typically developing longitudinal cohort

	16-month-olds			24-month-olds		
	N	*M*	*SD*	*N*	*M*	*SD*
Chronological Age	13	17.05	1.34	13	24.82	0.61
Mullen Scales of Early Learning						
Early Learning Composite	12	100.58	12.07	12	110.33	14.82
Verbal Composite	12	101.33	17.59	12	114.00	19.16
Non-Verbal Composite	12	109.67	8.91	12	110.08	16.94
Receptive Language						
Raw Score	12	17.92	2.88	13	27.31	2.66
T-Score	12	49.00	11.39	13	58.77	7.67
Expressive Language						
Raw Score	12	15.58	2.75	13	24.23	3.90
T-Score	12	47.25	9.19	13	53.54	10.16
MB-CDI Words and Gestures						
Phrase						
Number	13	20.69	8.30	-	-	-
Percentile	13	54.37	29.59	-	-	-
Understands						
Number	13	178.00	107.67	-	-	-
Percentile	13	45.08	34.13	-	-	-
Understands and Says						
Number	13	35.92	28.44	-	-	-
Percentile	13	43.77	24.83	-	-	-
MB-CDI Words and Sentences						
Words Produced						
Number	-	-	-	10	361.5	209.27
Percentile	-	-	-	10	59.80	32.91
Length of Three Longest Sentences						
Average	-	-	-	10	4.10	1.77
Percentile	-	-	-	10	52.80	25.83
Complexity						
Number	-	-	-	10	9.60	9.23
Percentile	-	-	-	10	65.40	18.95

Table 14. Eye tracking derived variables for the low-risk, typically developing, cross-sectional sample

	16 Months			24 Months		
	N	*M*	*SD*	*N*	*M*	*SD*
Proportion Fixation Duration						
Target	17	0.18	0.06	37	0.28	0.09
Category	17	0.38	0.08	37	0.36	0.05
Other Category	17	0.44	0.08	37	0.35	0.07
Proportion Fixation Count						
Target	17	0.16	0.04	37	0.24	0.07
Category	17	0.37	0.06	37	0.37	0.04
Other Category	17	0.47	0.07	37	0.39	0.07
Number of Runs						
Target Category	17	13.82	3.50	37	19.51	6.06
Other Category	17	12.06	4.34	37	13.62	4.09
Average Run Length						
Target Category	17	3.48	0.72	37	3.63	0.71
Other Category	17	3.54	0.83	37	2.80	0.45
Number of Runs Ratio	17	1.27	0.50	37	1.46	0.43
Average Run Length Ratio	17	1.03	0.28	37	1.31	0.27

Table 15. Eye tracking derived variables for the low-risk, typically developing longitudinal cohort

	16 Months			24 Months		
	N	*M*	*SD*	*N*	*M*	*SD*
Proportion Fixation Duration						
Target	13	0.18	0.05	13	0.27	0.08
Other Category	13	0.46	0.07	13	0.36	0.09
Proportion Fixation Count						
Target	13	0.16	0.04	13	0.23	0.07
Other Category	13	0.49	0.06	13	0.39	0.07
Number of Runs						
Target Category	13	14.00	4.00	13	20.54	6.86
Other Category	13	13.00	4.34	13	14.15	4.81
Average Run Length						
Target Category	13	3.58	0.75	13	3.37	0.50
Other Category	13	3.70	0.87	13	2.84	0.36
Number of Runs Ratio	13	1.17	0.45	13	1.48	0.43
Average Run Length Ratio	13	1.01	0.29	13	1.20	0.22

Table 16. Correlation matrix examining relationships between measures of the VAT and standardized measures of language for the low-risk, typically developing subsample at 16 months of age

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Proportion of Fixation Duration to Target	-						
(2) Proportion of Fixation Duration to Other-Category Objects	-.34	-					
(3) MSEL Early Learning Composite	.22	-.06	-				
(4) MSEL Receptive Language T-Score	.30	-.20	.74**	-			
(5) MB-CDI Phrase Percentile	.31	-.10	.38	.73**	-		
(6) MB-CDI Understands Percentile	.27	-.07	.52*	.77**	.90**	-	
(7) MB-CDI Understands and Says Percentile	.07	.06	.47	.52*	.57*	.51*	-

* correlation is significant at the .05 level (2-tailed)
** correlation is significant at the .01 level (2-tailed)

Table 17. Correlation matrix examining relationships between measures of the VAT and standardized measures of language for the low-risk, typically developing subsample at 24 months of age

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Proportion of Fixation Duration to Target	-						
(2) Proportion of Fixation Duration to Other-Category Objects	-.82**	-					
(3) MSEL Early Learning Composite	.25	-.45**	-				
(4) MSEL Receptive Language T-Score	.30	-.45**	.84**	-			
(5) MB-CDI Words Produced Percentile	-.03	.01	.46**	.53**	-		
(6) MB-CDI Length of Three Longest Sentences Percentile	-.28	.17	.40*	.41*	.72**	-	
(7) MB-CDI Complexity Percentile	-.22	.27	.34	.44*	.77**	.79**	-

* correlation is significant at the .05 level (2-tailed)
** correlation is significant at the .01 level (2-tailed)

Table 18. Correlation matrix examining relationships between the receptive language measure of the VAT and standardized measures of language for the low-risk, typically developing longitudinal cohort

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Proportion of Fixation Duration to Target at 16m	-											
(2) Proportion of Fixation Duration to Target at 24m	.61*	-										
(3) MSEL Early Learning Composite at 16	.46	.35	-									
(4) MSEL Early Learning Composite at 24m	.40	.39	.42	-								
(5) MSEL Receptive Language T-Score at 16m	.38	.28	.80**	.26	-							
(6) MSEL Receptive Language T-Score at 24m	.45	.46	.50	.86**	.38	-						
(7) MB-CDI Phrase Percentile	.05	-.06	.49	.38	.75**	.39	-					
(8) MB-CDI Understands Percentile	.25	.17	.65*	.37	.84**	.54	.88**	-				
(9) MB-CDI Understands and Says Percentile	-.25	-.26	.48	.09	.44	-.20	.51	.44	-			
(10) MB-CDI Words Produced Percentile	-.09	-.24	.28	.22	.57	.17	.93**	.73*	.68*	-		
(11) MB-CDI Length of Three Longest Sentences Percentile	-.14	-.30	.28	.15	.69*	.09	.87**	.74*	.66*	.90**	-	
(12) MB-CDI Complexity Percentile	-.14	-.41	.53	.08	.75*	-.02	.73*	.68*	.79**	.69*	.83**	-

* correlation is significant at the .05 level (2-tailed)
** correlation is significant at the .01 level (2-tailed)

Table 19. Correlation matrix examining relationships between variables of the VAT for the low-risk, typically developing cross-sectional cohort

	(1)	(2)	(3)	(4)
(1) Proportion Target Fixation Duration	-	-.82**	.70**	.33*
(2) Proportion Other-Category Fixation Duration	-.34	-	-.71**	-.49**
(3) Number of Runs Ratio	-.12	-.60**	-	.27
(4) MRL Ratio	.75	-.52*	-.05	-

* correlation is significant at the .05 level (2-tailed)
** correlation is significant at the .01 level (2-tailed)

Table 20. Eye tracking derived variables for the low-risk and high-risk cross-sectional sample

	Low-Risk			High-Risk		
	N	*M*	*SD*	*N*	*M*	*SD*
16 Months						
Proportion Fixation Duration						
Target	20	0.19	0.06	27	0.16	0.06
Category	20	0.38	0.08	27	0.39	0.09
Other Category	20	0.43	0.08	27	0.46	0.09
Proportion Fixation Count						
Target	20	0.17	0.05	27	0.15	0.05
Category	20	0.37	0.06	27	0.37	0.08
Other Category	20	0.46	0.08	27	0.47	0.07
Number of Runs						
Target Category	20	14.50	3.85	27	14.37	6.47
Other Category	20	12.00	4.10	26	13.46	5.67
Average Run Length						
Target Category	20	3.42	0.68	27	3.39	0.82
Other Category	20	3.44	0.81	26	3.26	0.77
Number of Runs Ratio	20	1.34	0.54	26	1.21	0.55
Average Run Length Ratio	20	1.03	0.26	26	1.08	0.26
24 Months						
Proportion Fixation Duration						
Target	42	0.28	0.09	22	0.26	0.09
Category	42	0.36	0.05	22	0.35	0.06
Other Category	42	0.36	0.07	22	0.40	0.09
Proportion Fixation Count						
Target	42	0.24	0.07	22	0.22	0.07
Category	42	0.37	0.05	22	0.36	0.06
Other Category	42	0.39	0.07	22	0.42	0.07
Number of Runs						
Target Category	42	19.52	5.99	22	17.55	4.67
Other Category	42	13.45	3.96	22	13.18	3.67
Average Run Length						
Target Category	42	3.60	0.70	22	3.36	0.63
Other Category	42	2.81	0.45	22	2.93	0.55
Number of Runs Ratio	42	1.48	0.43	22	1.39	0.39
Average Run Length Ratio	42	1.30	0.28	22	1.18	0.28

Table 21. Eye tracking derived variables for the low-risk and high-risk longitudinal cohort

	Low-Risk			High-Risk		
	N	*M*	*SD*	*N*	*M*	*SD*
16 Months						
Proportion Fixation Duration						
Target	14	0.18	0.5	9	0.16	0.02
Other Category	14	0.46	0.07	9	0.50	0.10
Proportion Fixation Count						
Target	14	0.17	0.03	9	0.17	0.06
Other Category	14	0.48	0.06	9	0.49	0.08
Number of Runs						
Target Category	14	14.36	4.07	9	12.56	5.00
Other Category	14	12.86	4.20	9	12.22	6.50
Average Run Length						
Target Category	14	3.53	0.74	9	3.32	0.91
Other Category	14	3.62	0.90	9	3.18	0.90
Number of Runs Ratio	14	1.21	0.46	9	1.26	0.65
Average Run Length Ratio	14	1.02	0.28	9	1.05	0.12
24 Months						
Proportion Fixation Duration						
Target	14	0.27	0.08	9	0.26	0.10
Other Category	14	0.36	0.08	9	0.40	0.06
Proportion Fixation Count						
Target	14	0.23	0.06	9	0.23	0.09
Other Category	14	0.39	0.07	9	0.43	0.07
Number of Runs						
Target Category	14	20.79	6.66	9	15.56	5.22
Other Category	14	14.36	4.68	9	10.89	3.02
Average Run Length						
Target Category	14	3.36	0.48	9	3.29	0.41
Other Category	14	2.85	0.35	9	3.09	0.77
Number of Runs Ratio	14	1.48	0.41	9	1.46	0.42
Average Run Length Ratio	14	1.19	0.21	9	1.11	0.27

Table 22. Correlation matrix examining relationships between variables of the VAT for the low-risk cross-sectional sample

	(1)	(2)	(3)	(4)
(1) Proportion Target Fixation Duration	-	-.79**	.68**	.32*
(2) Proportion Other-Category Fixation Duration	-.46*	-	-.69**	-.46**
(3) Number of Runs Ratio	.11	-.69**	-	.35
(4) MRL Ratio	.07	-.47*	-.03	-

* correlation is significant at the .05 level (2-tailed)
** correlation is significant at the .01 level (2-tailed)

Note: Correlations for the 16-month-old group are displayed on the bottom of the chart and correlations for the 24-month-old group are displayed on the top of the chart.

Table 23. Correlation matrix examining relationships between variables of the VAT for the high-risk cross-sectional sample

	(1)	(2)	(3)	(4)
(1) Proportion Target Fixation Duration	-	-.77**	.44*	.44*
(2) Proportion Other-Category Fixation Duration	-.37	-	-.41	-.68**
(3) Number of Runs Ratio	.38	-.45*	-	.09
(4) MRL Ratio	-.02	-.21	-.38	-

* correlation is significant at the .05 level (2-tailed)
** correlation is significant at the .01 level (2-tailed)

Note: Correlations for the 16-month-old group are displayed on the bottom of the chart and correlations for the 24-month-old group are displayed on the top of the chart.

Table 24. Eye tracking derived variables for the cross-sectional sample outcome groups

	TD			ASD			NT		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
16 Months									
Proportion Fixation Duration									
Target	33	0.17	0.06	5	0.17	0.07	6	0.16	0.03
Category	33	0.38	0.07	5	0.37	0.08	6	0.38	0.14
Other Category	33	0.45	0.08	5	0.45	0.05	6	0.50	0.14
Proportion Fixation Count									
Target	33	0.16	0.04	5	0.18	0.07	6	0.18	0.05
Category	33	0.37	0.06	5	0.37	0.06	6	0.36	0.15
Other Category	33	0.47	0.06	5	0.45	0.02	6	0.46	0.12
Number of Runs									
Target Category	33	13.82	4.98	5	17.00	8.22	6	13.33	5.85
Other Category	33	12.88	5.06	5	12.20	6.57	6	12.17	4.88
Average Run Length									
Target Category	32	3.44	0.79	5	2.74	0.36	6	3.62	0.72
Other Category	32	3.40	0.81	5	3.08	0.94	6	3.17	0.77
Number of Runs Ratio	32	1.12	0.52	5	1.60	0.69	6	1..15	0.41
Average Run Length Ratio	32	1.06	0.28	5	0.93	0.18	6	1.17	0.21
24 Months									
Proportion Fixation Duration									
Target	51	0.28	0.09	5	0.17	0.04	7	0.30	0.10
Category	51	0.36	0.06	5	0.37	0.07	7	0.36	0.05
Other Category	51	0.36	0.07	5	0.46	0.10	7	0.34	0.07
Proportion Fixation Count									
Target	51	0.24	0.07	5	0.16	0.02	7	0.25	0.07
Category	51	0.37	0.05	5	0.38	0.06	7	0.39	0.04

Other Category	51	0.40	0.07	5	0.46	0.06	7	0.36	0.05
Number of Runs									
Target Category	51	18.92	6.00	5	17.00	1.73	7	20.57	3.86
Other Category	51	13.45	4.04	5	15.20	2.86	7	11.86	2.41
Average Run Length									
Target Category	51	3.60	0.69	5	2.87	0.53	7	3.52	0.49
Other Category	51	2.82	0.42	5	2.71	0.42	7	3.20	0.85
Number of Runs Ratio	51	1.44	0.42	5	1.14	0.18	7	1.76	0.31
Average Run Length Ratio	51	1.29	0.27	5	1.09	0.30	7	1.17	0.36

Table 25. Eye tracking derived variables for the longitudinal cohort outcome groups

	TD			ASD			NT		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
16 Months									
Proportion Fixation Duration									
Target	19	0.17	0.04	1	0.19	-	3	0.17	0.02
Category	19								
Other Category	19	0.46	0.07	1	0.52	-	3	0.53	0.17
Proportion Fixation Count									
Target	19	0.16	0.03	1	0.24	-	3	0.20	0.06
Category	19								
Other Category	19	0.48	0.06	1	0.45	-	3	0.53	0.11
Number of Runs									
Target Category	19	14.05	4.06	1	7.00	-	3	13.33	6.66
Other Category	19	13.11	4.97	1	3.00	-	3	12.67	3.79
Average Run Length									
Target Category	19	3.54	0.80	1	2.24	-	3	3.25	0.55
Other Category	19	3.58	0.84	1	2.00	-	3	3.08	1.07
Number of Runs Ratio	19	1.20	0.48	1	2.33	-	3	1.07	0.59
Average Run Length Ratio	19	1.02	0.24	1	1.12	-	3	1.09	0.18
24 Months									
Proportion Fixation Duration									
Target	19	0.27	0.09	1	0.22	-	3	0.29	0.12
Category	19								
Other Category	19	0.38	0.08	1	0.42	-	3	0.37	0.10
Proportion Fixation Count									
Target	19	0.24	0.07	1	0.18	-	3	0.24	0.08
Category	19								
Other Category	19	0.39	0.07	1	0.41	-	3	0.39	0.06

Number of Runs									
Target Category	19	18.47	7.10	1	17.00	-	3	21.00	3.00
Other Category	19	12.89	4.68	1	15.00	-	3	13.00	3.61
Average Run Length									
Target Category	19	3.36	0.48	1	2.98	-	3	3.25	0.20
Other Category	19	2.86	0.35	1	2.43	-	3	3.65	1.12
Number of Runs Ratio	19	1.46	0.43	1	1.13	-	3	1.65	0.21
Average Run Length Ratio	19	1.19	0.23	1	1.23	-	3	0.93	0.20

Table 26. Correlation matrix examining relationships between variables of the VAT for the typically developing cross-sectional sample

	(1)	(2)	(3)	(4)
(1) Proportion Target Fixation Duration	-	-.77**	.60**	.37**
(2) Proportion Other-Category Fixation Duration	-.45*	-	-.62**	-.56**
(3) Number of Runs Ratio	.12	-.62**	-	.19
(4) MRL Ratio	.02	-.35	-.25	-

* correlation is significant at the .05 level (2-tailed)
** correlation is significant at the .01 level (2-tailed)

Note: Correlations for the 16-month-old group are displayed on the bottom of the chart and correlations for the 24-month-old group are displayed on the top of the chart.

Table 27. Correlation matrix examining relationships between variables of the VAT for the non-typically developing cross-sectional sample

	(1)	(2)	(3)	(4)
(1) Proportion Target Fixation Duration	-	-.89**	.43	.31
(2) Proportion Other-Category Fixation Duration	-.25	-	-.27	-.37
(3) Number of Runs Ratio	.42	-.60	-	-.18
(4) MRL Ratio	.05	-.66	.03	-

* correlation is significant at the .05 level (2-tailed)
** correlation is significant at the .01 level (2-tailed)

Note: Correlations for the 16-month-old group are displayed on the bottom of the chart and correlations for the 24-month-old group are displayed on the top of the chart.

Table 28. Correlation matrix examining relationships between variables of the VAT for the ASD cross-sectional sample

	(1)	(2)	(3)	(4)
(1) Proportion Target Fixation Duration	-	-.75	.37	.46
(2) Proportion Other-Category Fixation Duration	-.17	-	-.29	-.87
(3) Number of Runs Ratio	.35	.02	-	.11
(4) MRL Ratio	.15	.90*	-.06	-

* correlation is significant at the .05 level (2-tailed)
** correlation is significant at the .01 level (2-tailed)

Note: Correlations for the 16-month-old group are displayed on the bottom of the chart and correlations for the 24-month-old group are displayed on the top of the chart.

6.0 FIGURES

Figure 1. Example of visual stimulus projected on screen. The eight color illustrations depict items listed in the Mullen Scales of Early Learning (Mullen, 1995) and the MacArthur-Bates Communicative Development Inventories (Fenson et al., 1993).

4	1	0	4	8
4	2	3	2	0
10	1	1	5	6
8	3	3	7	0

Figure 2. Schematic representing the number of times the target object appeared in each location within the 5 (width) x 4 (height) stimulus grid.

Order	A	B	C	D	E	F
			Target Word			
	Animal/Food	*Vehicle/Utensil*	*Furniture/Animal*	*Clothing/Vehicle*	*Food/Furniture*	*Utensil/Clothing*
1	Dog	Car	Table	Hat	Apple	Spoon
2	Horse	Train	Door	Pants	Banana	Bottle
3	Bird	Airplane	Chair	Shoe	Cookie	Cup
4	Cat	Truck	Bed	Shirt	Bread	Plate

	Order of Stimuli Presentation		
Stimulus Set	**Orders Received**	**First Six Trials**	**Second Six Trials**
1	1, 2	A B C D E F	A B C D E F
2	1, 3	B C D E F A	B C D E F A
3	1, 4	C D E F A B	C D E F A B
4	2, 3	D E F A B C	D E F A B C
5	2, 4	E F A B C D	E F A B C D
6	3, 4	F A B C D E	F A B C D E

Figure 3. System used to counterbalance target word presentation and category contrasts across the six stimulus sets.

Figure 4. Example areas of interest (AOIs) depicting the 190 x 190pixel region defined for each item as well as the entirety of the projection screen.

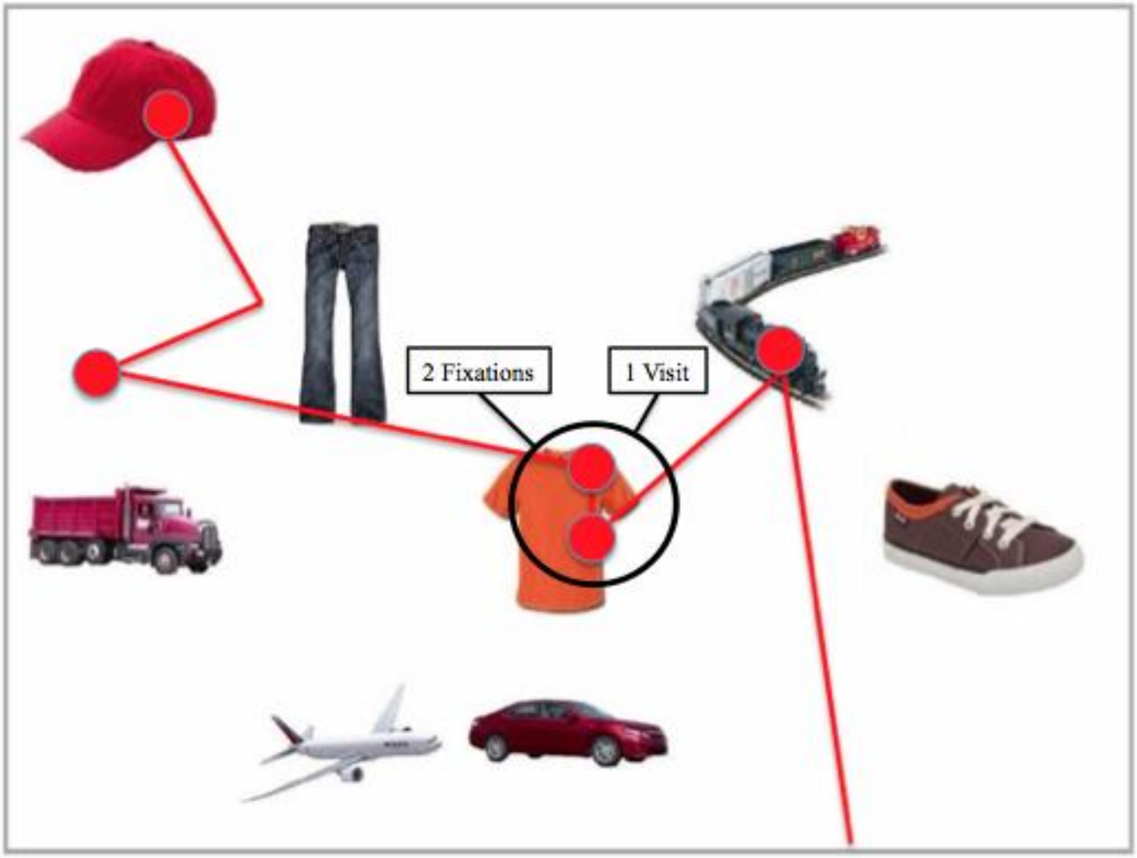

Figure 5. Example scanpath demonstrating the difference in criteria for a fixation and a visit.

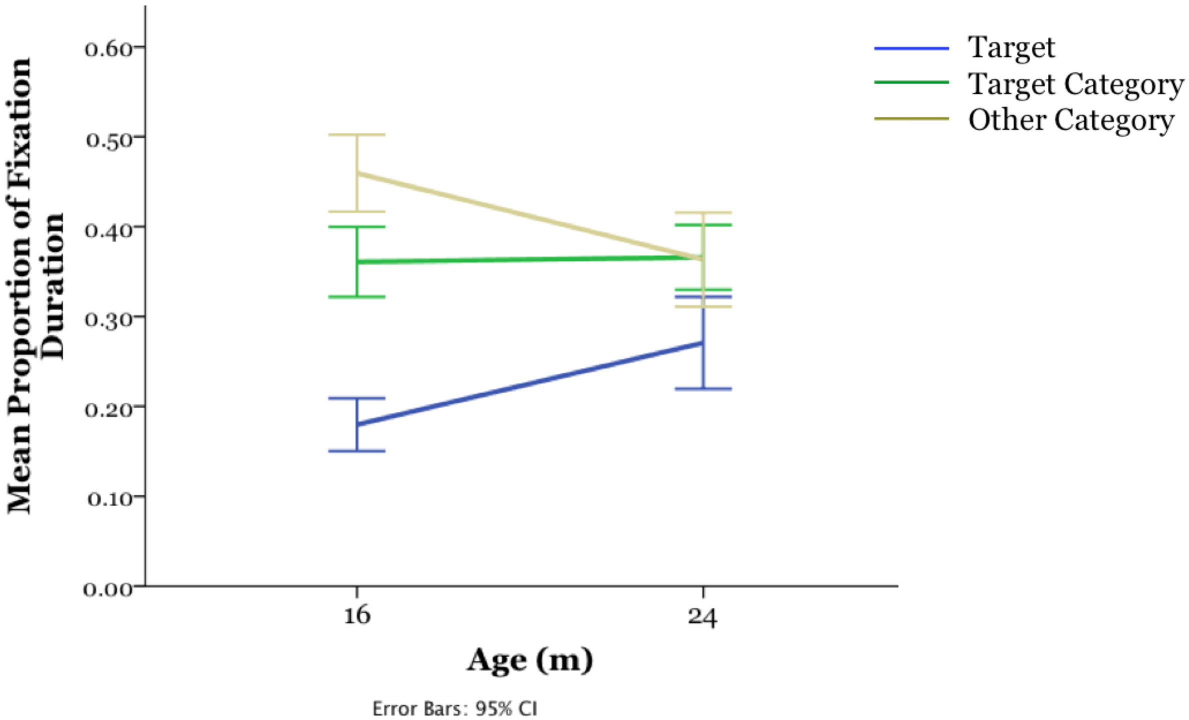

Figure 6. Developmental trajectories for the proportion of target fixation duration, proportion of target-category fixation duration, proportion of other-category fixation duration for the longitudinal cohort of the low-risk, typically developing subsample.

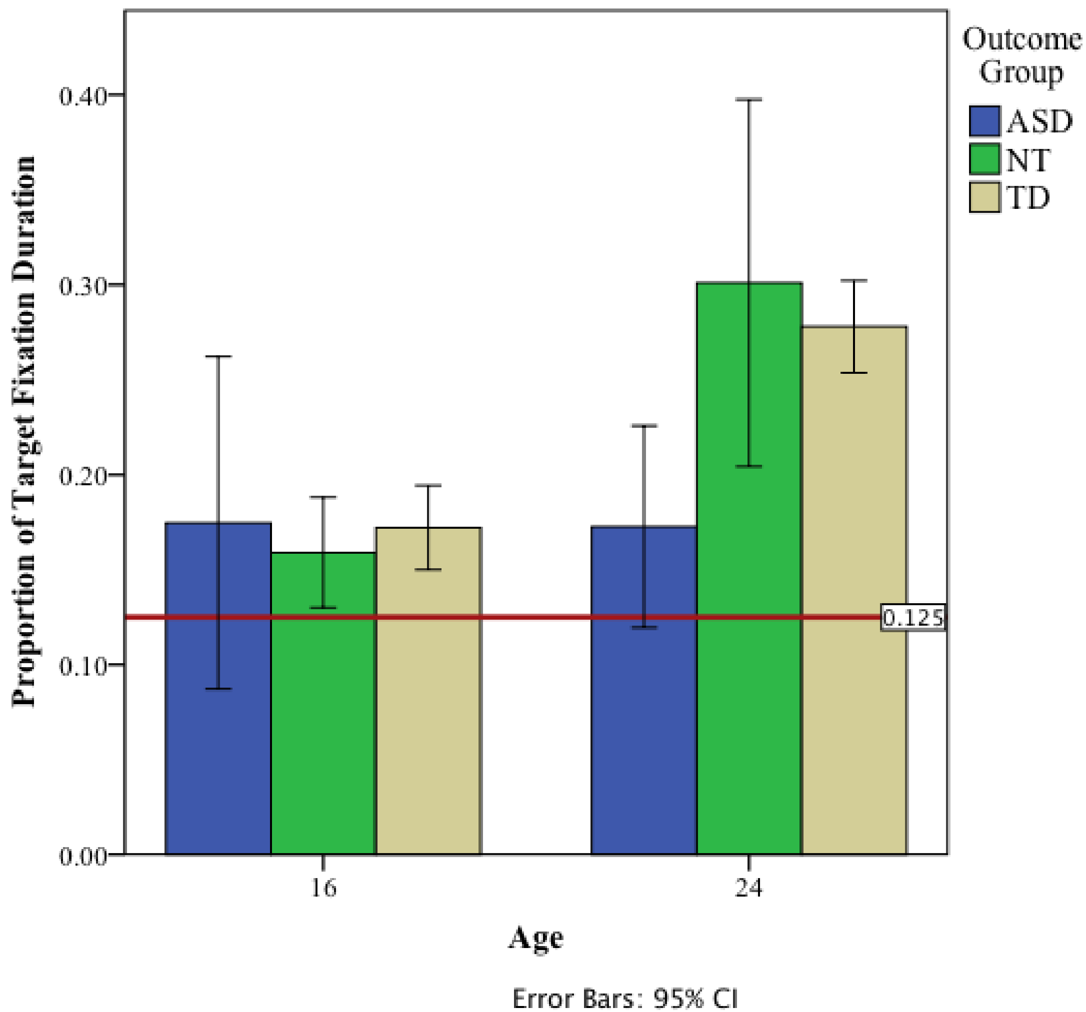

Figure 7. The proportion of time spent looking to the target object by outcome groups at 16 and 24 months of age.

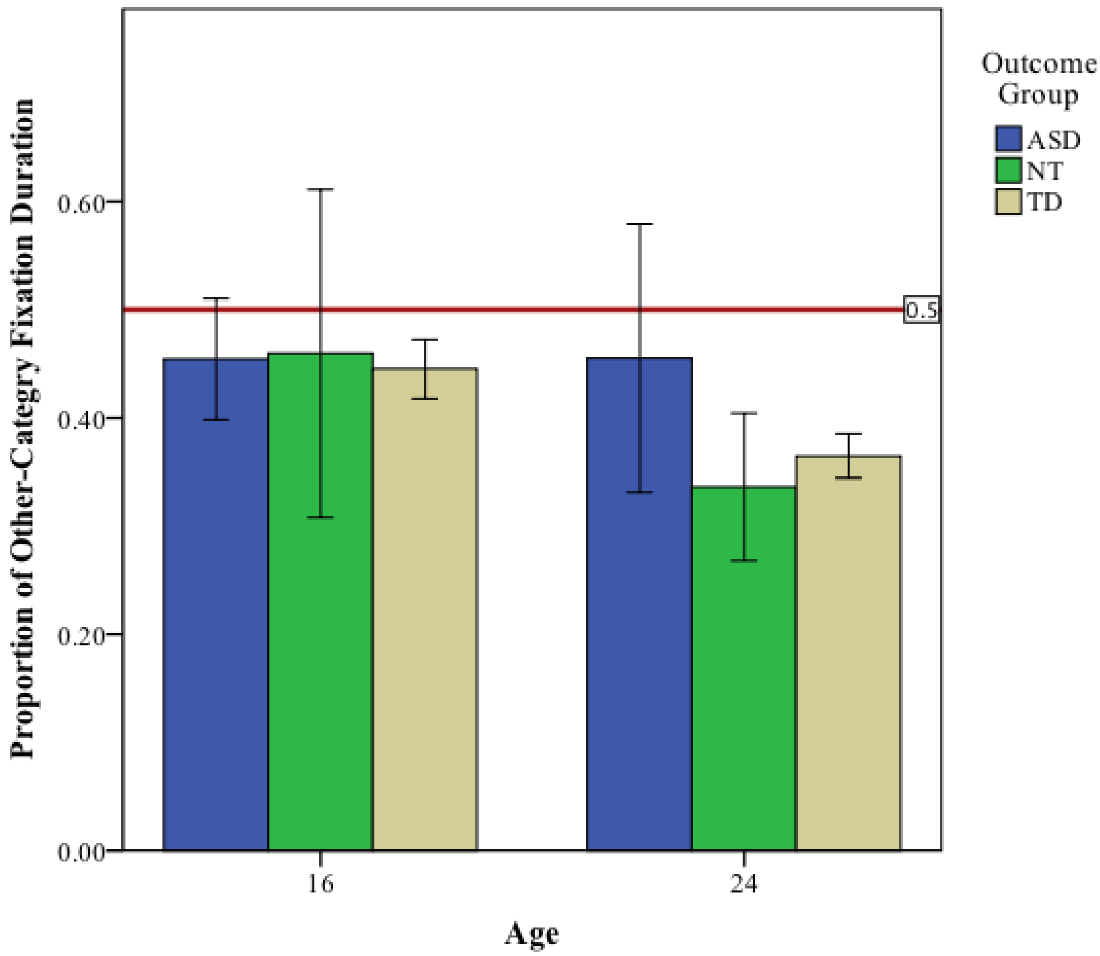

Figure 8. The proportion of time spent looking to the other-category objects by outcome groups at 16 and 24 months of age.

BIBLIOGRAPHY

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., ... & Durkin, M. S. (2018). Prevalence of autism spectrum disorder among children aged 8 years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *MMWR Surveillance Summaries*, *67*(6), 1.

Bavin, E. L., Kidd, E., Prendergast, L., Baker, E., Dissanayake, C., & Prior, M. (2014). Severity of autism is related to children's language processing. *Autism Research*, *7*(6), 687-694.

Bayley, N. (1993). *Bayley scales of infant development: Manual*. Psychological Corporation.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253-3258.

Bergelson, E., & Swingley, D. (2015). Early word comprehension in infants: Replication and extension. *Language Learning and Development*, *11*(4), 369-380.

Brady, N. C., Anderson, C. J., Hahn, L. J., Obermeier, S. M., & Kapa, L. L. (2014). Eye tracking as a measure of receptive vocabulary in children with autism spectrum disorders. *Augmentative and Alternative Communication*, *30*(2), 147-159.

Bruckner, C., Yoder, P., Stone, W., & Saylor, M. (2007). Construct validity of the MB-CDI-I receptive vocabulary scale can be improved: Differential item functioning between toddlers with autism spectrum disorders and typically developing infants. *Journal of Speech, Language, and Hearing Research*, *50*(6), 1631-1638.

Charman, T., Baron- Cohen, S., Swettenham, J., Baird, G., Drew, A., & Cox, A. (2003). Predicting language outcome in infants with autism and pervasive developmental disorder. *International Journal of Language & Communication Disorders*, *38*(3), 265-285.

Charman, T., Young, G. S., Brian, J., Carter, A., Carver, L. J., Chawarska, K., ... & Hertz- Picciotto, I. (2017). Non- ASD outcomes at 36 months in siblings at familial risk for autism spectrum disorder (ASD): A baby siblings research consortium (BSRC) study. *Autism Research*, *10*(1), 169-178.

Chita-Tegmark, M., Arunachalam, S., Nelson, C. A., & Tager-Flusberg, H. (2015). Eye-tracking measurements of language processing: developmental differences in children at high risk for ASD. *Journal of autism and developmental disorders*, *45*(10), 3327-3338.

Coonrod, E. E., & Stone, W. L. (2004). Early concerns of parents of children with autistic and nonautistic disorders. *Infants & Young Children*, *17*(3), 258-268.

Dunn, L. M., & Dunn, L. M. (1997). *PPVT-III: Peabody picture vocabulary test*. American Guidance Service.

Fantz, R. L. (1958). Pattern vision in young infants. *The psychological record*, *8*(2), 43-47.

Ferguson, B., & Waxman, S. (2017). Linking language and categorization in infancy. *Journal of child language*, *44*(3), 527-552.

Fenson, L., Bates, E., Dale, P. S., Hartung, J. P., Pethick, S. J., Reilly, J. S., ... & Thal, D. J. (1996). *The MacArthur communicative development inventories*. Singular publishing Gruop, Incorporated.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., ... & Stiles, J. (1994). Variability in early communicative development. *Monographs of the society for research in child development*, i-185.

Fernald, A., Pinto, J. P., Swingley, D., Weinbergy, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, *9*(3), 228-231.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language. *Developmental psycholinguistics: On-line methods in children's language processing*, *44*, 97.

Gastgeb, H. Z., Dundas, E. M., Minshew, N. J., & Strauss, M. S. (2012). Category formation in autism: can individuals with autism form categories and prototypes of dot patterns?. *Journal of autism and developmental disorders*, *42*(8), 1694-1704.

Gastgeb, H. Z., Rump, K. M., Best, C. A., Minshew, N. J., & Strauss, M. S. (2009). Prototype formation in autism: can individuals with autism abstract facial prototypes?. *Autism Research*, *2*(5), 279-284.

Gastgeb, H. Z., & Strauss, M. S. (2012). Categorization in ASD: The role of typicality and development. *Perspectives on language learning and education*, *19*(2), 66-74.

Gastgeb, H. Z., Strauss, M. S., & Minshew, N. J. (2006). Do individuals with autism process categories differently? The effect of typicality and development. *Child development*, *77*(6), 1717-1729.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of child language*, *14*(1), 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.

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: the case for a developmental lexical principles framework [*]. *Journal of child language*, *21*(1), 125-155.

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of child language*, *27*(3), 689-705.

Houston-Price, C., Mather, E., & Sakkalou, E. (2007). Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *Journal of Child Language*, *34*(4), 701-724.

Jones, E. J., Gliga, T., Bedford, R., Charman, T., & Johnson, M. H. (2014). Developmental pathways to autism: a review of prospective studies of infants at risk. *Neuroscience & Biobehavioral Reviews*, *39*, 1-33.

Kasari, C., Brady, N., Lord, C., & Tager ‐ Flusberg, H. (2013). Assessing the minimally verbal school ‐ aged child with autism spectrum disorder. *Autism Research*, *6*(6), 479-493.

Kwok, E. Y., Brown, H. M., Smyth, R. E., & Cardy, J. O. (2015). Meta-analysis of receptive and expressive language skills in autism spectrum disorder. *Research in Autism Spectrum Disorders*, *9*, 202-222.

Law, J., & Roy, P. (2008). Parental report of infant language skills: A review of the development and application of the Communicative Development Inventories. *Child and Adolescent Mental Health*, *13*(4), 198-206.

Landa, R., & Garrett- Mayer, E. (2006). Development in infants with autism spectrum disorders: a prospective study. *Journal of Child Psychology and Psychiatry*, *47*(6), 629-638.

Lord, C. (1995). Follow- up of two- year- olds referred for possible autism. *Journal of child psychology and psychiatry*, *36*(8), 1365-1382.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2001). *Autism diagnostic observation schedule (ADOS): Manual*. WPS.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of autism and developmental disorders*, *24*(5), 659-685.

Luyster, R. J., Kadlec, M. B., Carter, A., & Tager-Flusberg, H. (2008). Language assessment and development in toddlers with autism spectrum disorders. *Journal of autism and developmental disorders*, *38*(8), 1426-1438.

Luyster, R., Lopez, K., & Lord, C. (2007). Characterizing communicative development in children referred for autism spectrum disorders using the MacArthur-Bates Communicative Development Inventory (MB-CDI). *Journal of child language*, *34*(3), 623-654.

Luyster, R. J., Seery, A., Talbott, M. R., & Tager- Flusberg, H. (2011). Identifying Early- Risk Markers and Developmental Trajectories for Language Impairment in Neurodevelopmental Disorders. *Developmental disabilities research reviews*, *17*(2), 151-159.

Mareschal, D., & Quinn, P. C. (2001). Categorization in infancy. *Trends in cognitive sciences*, *5*(10), 443-450.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive psychology*, *20*(2), 121-157.

Mandler, J. M., & Bauer, P. J. (1988). The cradle of categorization: Is the basic level basic?. *Cognitive development*, *3*(3), 247-264.

Mandler, J. M., Bauer, P. J., & McDonough, L. (1991). Separating the sheep from the goats: Differentiating global categories. *Cognitive Psychology*, *23*(2), 263-298.

Mandler, J. M., Fivush, R., & Reznick, J. S. (1987). The development of contextual categories. *Cognitive Development*, *2*(4), 339-354.

Mervis, C. B., & Bertrand, J. (1994). Acquisition of the novel name–nameless category (N3C) principle. *Child development*, *65*(6), 1646-1662.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual review of psychology*, *32*(1), 89-115.

Messinger, D. S., Young, G. S., Webb, S. J., Ozonoff, S., Bryson, S. E., Carter, A., ... & Dobkins, K. (2015). Early sex differences are not autism-specific: a Baby Siblings Research Consortium (BSRC) study. *Molecular autism*, *6*(1), 32.

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. *Journal of Developmental & Behavioral Pediatrics*, *27*(2), S69-S78.

Mullen, E. M. (1995). *Mullen scales of early learning* (pp. 58-64). Circle Pines, MN: AGS.

Newell, L. C., Best, C. A., Gastgeb, H., Rump, K. M., Strauss, M. S., Oakes, L. M., ... & Rakison, D. H. (2010). The development of categorization and facial knowledge: Implications for the study of autism. *Infant perception and cognition: Recent advances, emerging theories, and future directions*, 223-260.

Olsson, P. (2007). Real-time and offline filters for eye tracking. Msc thesis, KTH Royal Institute of Technology, April 2007.

Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., ... & Hutman, T. (2011). Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium study. *Pediatrics*, *128*(3), e488-e495.

Pickles, A., Simonoff, E., Conti- Ramsden, G., Falcaro, M., Simkin, Z., Charman, T., ... & Baird, G. (2009). Loss of language in early development of autism and specific language impairment. *Journal of Child Psychology and Psychiatry*, *50*(7), 843-852.

Plesa Skwerer, D., Jordan, S. E., Brukilacchio, B. H., & Tager-Flusberg, H. (2016). Comparing methods for assessing receptive language skills in minimally verbal children and adolescents with autism spectrum disorders. *Autism*, *20*(5), 591-604.

Quinn, P. C., & Johnson, M. H. (2000). Global- Before- Basic Object Categorization in Connectionist Networks and 2- Month- Old Infants. *Infancy*, *1*(1), 31-46.

Ricciuti, H. N. (1965). Object grouping and selective ordering behavior in infants 12 to 24 months old. *Merrill-Palmer Quarterly of Behavior and Development*, *11*(2), 129-148.

Rosch, E. (1978). Principles of categorization. Cognition and categorization, ed. by Eleanor Rosch & Barbara B. Lloyd, 27-48.

Starkey, D. (1981). The origins of concept formation: Object sorting and object preference in early infancy. *Child Development*, 489-497.

Sparrow, S. S., Balla, D. A., Cicchetti, D. V., Harrison, P. L., & Doll, E. A. (1984). Vineland adaptive behavior scales.

Tager-Flusberg, H. (1985a). Basic level and superordinate level categorization by autistic, mentally retarded, and normal children. *Journal of Experimental Child Psychology*, *40*(3), 450-469.

Tager-Flusberg, H. (1985b). The conceptual basis for referential word meaning in children with autism. *Child Development*, 1167-1178.

Tager-Flusberg, H. (1999). A psychological approach to understanding the social and language impairments in autism. *International review of psychiatry*, *11*(4), 325-334.

Tager-Flusberg, H. (2016). Risk factors associated with language in autism spectrum disorder: Clues to underlying mechanisms. *Journal of Speech, Language, and Hearing Research*, *59*(1), 143-154.

Tager- Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism. *Handbook of Autism and Pervasive Developmental Disorders, Volume 1, Third Edition*, 335-364.

Tomasello, M., & Mervis, C. B. (1994). The instrument is great, but measuring comprehension is still a problem. *Monographs of the Society for Research in child Development*, *59*(5), 174-179.

Ungerer, Judy A., and Marian Sigman. "Categorization skills and receptive language development in autistic children." *Journal of Autism and Developmental Disorders* 17, no. 1 (1987): 3-16.

Vitrano, J. L. (2015). *Object categorization in children with autism spectrum disorder (ASD)*. City University of New York.

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental science*, *10*(2), 159-164.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12-to 13-month-old infants. *Cognitive psychology*, *29*(3), 257-302.

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 and developmental disorders*, *40*(10), 1259-1273.

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *International journal of developmental neuroscience*, *23*(2), 143-152.