Responsiveness to Joint Attention in Autism: Predictive Characteristics and Concurrent Mechanisms
UNIVERSITY OF CALIFORNIA

Los Angeles

Responsiveness to Joint Attention in Autism: Predictive Characteristics and Concurrent Mechanisms

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

by

Kristen Marie Gillespie-Lynch

2012
ABSTRACT OF THE DISSERTATION

Responsiveness to Joint Attention in Autism: Predictive Characteristics and Concurrent Mechanisms

by

Kristen Marie Gillespie-Lynch

University of California, Los Angeles, 2012

Professor Scott P. Johnson, Chair

What are the developmental implications and underlying mechanisms of atypical responsiveness to joint attention (RJA) in autism? A longitudinal investigation of relations between individual differences in early childhood RJA and adult outcomes revealed associations between early joint attention and adult adaptive skills, symptoms and social functioning. Many of these associations were attributable to relations between RJA and subsequent changes in cognitive and linguistic skills which in turn predicted adult skills and symptoms. A second longitudinal study assessed relations between an eye-tracking measure of social responsiveness and gaze following administered in infancy and subsequent symptoms and diagnoses of autism. Social responsiveness, specifically interest in dyadic cues, but not gaze following was associated with subsequent autism diagnoses. A third study used eye-tracking and in-person assessments to examine three potential mechanisms underlying atypical gaze following in autism: impaired reflexive gaze cueing, difficulty integrating gaze and affect or reduced recognition of the referential significance of gaze. Reduced low- and high-level gaze following were associated with autism while difficulties with reference were related to developmental level rather than
autism. No evidence of gaze and emotion integration was observed regardless of diagnosis. Thus, reduced RJA in autism may arise from low-level atypicalities in gaze following while impaired integration of gaze and affect and reduced understanding of reference may emerge with development. Together these studies imply that individual differences in dyadic and triadic aspects of social attention probably arise from low level mechanisms and have long-term implications for individuals with or at-risk for autism.
The dissertation of Kristen Marie Gillespie-Lynch is approved.

Mirella Dapretto

Patricia M. Greenfield

Connie Kasari

Ladan Shams

Nim Delafield

Scott P. Johnson, Committee Chair

University of California, Los Angeles

2012
Dedication Page

This work is dedicated to Doris Lynch (for sharing a love of words with me since before I was born), to Thom Gillespie (for sharing a love of mysteries and the stories we tell to make sense of them), to Cody Gillespie-Lynch (for sharing his love of lost causes), to David Smith (for being here to talk about it all and making me beautiful stimuli), to the children who inspired me with their joys and sorrows, and to Marian Sigman whose lifework was my North Star flickering between a precise map of previously uncharted aspects of autism and the exceptions to every rule, your wit and passion are missed.

Taos Sky by Doris Lynch

How do you describe this vault of longing?

Sky that makes you want to lie in a scrubby field all day and pay attention. We humans weren’t made for such devoted attention.

For distraction: rasps of cottonwood leaves, magpie’s swoop and gossipy chatter,

and the long-haired builder hammering. Even the ponderosa pine pulls as

It angles tipsily and its needles patch feathery green islands across this canvas of sky. What more to want than this cerulean Heaven?

Only to focus each moment on enduring blue, to watch, to observe, to breathe in this vastness until miniscule eyes erupt and weep silver.
Table of Contents

Abstract.........................................................................................................................ii

Dedication Page............................................................................................................v

Table of Contents..........................................................................................................vi

Acknowledgements......................................................................................................vii

Preface.........................................................................................................................viii

Vita...............................................................................................................................ix

Introduction..................................................................................................................1

Early Childhood Predictors of the Social Competence of Adults with Autism..............4

Eye-Tracking Gaze Following in Infants at Risk for Autism..........................................31

Atypical Gaze Following in Autism: A Comparison of Three Potential Mechanisms......57

References....................................................................................................................88
Acknowledgements

Thank you to the children and their parents who taught me more than school could teach. Thank you to my committee members who each provided something unique and valuable. Mirella Dapretto, I have always been inspired by the vibrant originality of your research (and you’ve provided excellent advice on experimental design). Patricia Greenfield, I will never forget your generous mentorship. You always stepped in and provided guidance when I most needed it. Your lab was a warm and thought provoking intellectual home throughout my years at UCLA. Connie Kasari, I hope to find a way to bring the things I’ve learned here back to help people in the way your research does (and without your recruitment ideas for the third study, I doubt I would have finished it). Ladan Shams, your curiosity and sensible insights were much appreciated (such as that I should first assess emotion understanding if I want to assess emotion integration). Nim Tottenham, your thoughtful questions have taught me a great deal about what it means to be a valuable member of an intellectual community (and your hands on guidance in higher-level uses of excel and SPSS was invaluable). Scott Johnson, without you I would probably not have been able to focus on autism at UCLA (and I learned so much about clever experimental design from you). Ted Hutman, you were not on my committee due to a technicality but definitely should have been as you provided consistent mentorship, intellectual guidance, and dialogue about the exciting murkiness of autism research throughout graduate school. Thank you to my fellow researchers who provided a sense that it was all worth it: Adriana Manago, Rachel Brezis, Yalda T. Uhls, Emmy Goldknopf, Monika Abels, Goldie Salimkhkan, Angie Guan, Heejung Park, Lauren Sherman, Steven Kapp, Stephanie Marshall, Leigh Sepeta, Mari Davies, Mithi del Rosario, Lovella Gomez, Brigid McCarthy, Zoe Samson, Rebecca Elias, Jane Lee, Nancy Le, Jen Loa, Kia Dela Cruz, Devi Beck-Pancer, Yunping Feng and Christine Park.
Preface

My interest in autism arose from my experiences working with a series of fascinating youth with autism. Ten years ago, I started working at a group home for adolescents with severe disabilities. Although I didn’t know it when I first started working there, the organization was known as “the end of the road” or the place to put people who had nowhere else to go. One of the three boys in the group home was a nonverbal 15 year old with autism. Most of the time he was gentle and affectionate and liked to take caregivers by the hand to blow bubbles with him while watching Tim Burton movies. But sometimes, in a flash, he would grab people by the hair and yank them about or bite them, resulting in serious injuries. It seemed clear that these sudden outbursts were related to difficulty communicating but management focused on controlling behaviors rather than preventing them by helping him communicate. I helped him learn a few signs and started a picture schedule system for him. Two of the boys became attached to me, probably because of my attempts to help them communicate, and one day on an ill-conceived trip to the arcade a teenager with brain trauma bit me, and broke the fingers of another caregiver, because I was going to go into the arcade with the boy with autism rather than him. A co-worker offered me a position as an in-home caregiver for seven year old identical female twins with autism and I left the boys behind. As a caregiver with a college degree and no previous experience with disabilities, I did not feel that I had the resources to exert the systemic change that needed to occur at the “end of the road.”

One of the twins had been considered less severely autistic when first diagnosed but she had not learned to speak. Her sister loved to sing and had learned to speak in short phrases through song. Over the course of three years working with the twins, I helped them develop language through sign, pictures and by tying communication to activities they enjoyed. However,
the therapist who set up their treatment plans had limited contact with the twins and little training in autism. For example, when one of the twins took to pretending to be blind, the therapist thought it was because of the tile patterns on the floor. It began to seem like inadequately trained supervisors were a systemic issue impacting the lives of children with severe forms of autism. I decided that I could have more of an impact by getting further training. Although the twins progressed a great deal while I and another caregiver who was very good with them were working with them, they had lost most of their language when I visited them years later. I almost didn’t get to see them because the less verbal one was in the hospital after breaking a window.

While getting a master’s degree, I took courses in applied behavior analysis (ABA) and worked as a teacher’s assistant at an ABA school for children with autism. While I was initially drawn to the logic of ABA (it contrasted pleasantly with the lack of logic I had often encountered while working as a caregiver), in practice it only worked for a subset of the children I worked with at the school. While some blossomed, others struggled with the same tasks (i.e. labeling the first three letters of the alphabet in exchange for junk food reinforcements) for years. While working there, I tried to implement intervention approaches that used the children’s interests to support their social-communicative development.

Although these experiences were challenging and sometimes pretty heart breaking, the children were often so full of joy and I admired the singularity of their interests. In many ways they seemed wiser than the rest of us with our passions so watered down by adapting to other people’s perspectives. My primary goal as a caregiver was always to help children communicate by making communication relevant to them. The line of research in this dissertation arises from these experiences and the desire to increase our understanding of the roots of communication in autism, including how these roots grow into the future.
Vita

EDUCATION

2006 M.A. General Psychology, Queens College, City University of New York
2001 B.A. General Studies, Indiana University

GRANTS AND AWARDS

Center for Culture, Brain and Development Stipend, University of California, Los Angeles (2009-2012)
Summer Research Mentorship Award, University of California, Los Angeles (2008-2009)
Queens College Exemplary Scholarship Commendation (2006)
National Merit Scholar

PUBLICATIONS


SELECTED PRESENTATIONS


Introduction

Responsiveness to joint attention (RJA), defined as gaze or point following, begins to emerge very early in human development (D'Entremont, Hains & Muir, 1997, Farroni, Massaccesi, Pividori, & Johnson, 2004; Hood, Willen, & Driver, 1998; Scaife & Bruner, 1975) and develops gradually from birth to around 18 months of age (Butler, Caron & Brooks, 2009; Butterworth & Jarrett, 1991; Deák, Flom, & Pick, 2000). It is apparent across a range of species including ravens (Schloegl, Kotrschal, & Bugnyar, 2007), monkeys (Teufel, Gutmann, Pirow, & Fischer, 2010), apes (Tomasello, Hare, & Agnetta, 1999) and dolphins (Pack & Herman, 2004). Though it is rooted deep in ontogeny and phylogeny, it is not always observed, particularly among children with autism (e.g. Dawson et al., 2004; Loveland & Landry, 1986; Sigman & Ruskin, 1999). Indeed, reduced RJA in infancy is an early predictor of autism (Rozga et al., 2011; Landa, Holman, & Garrett-Mayer, 2007, Sullivan et al., 2007; Yoder, Stone, Walden, & Malesa, 2009). Among individuals with ASCs (autism spectrum conditions), individual differences in RJA are related to subsequent cognitive (Sigman & Ruskin, 1999), social-communicative (Mundy, Sigman, & Kasari, 1990; Sigman & Ruskin, 1999; Sigman & McGovern, 2005) and adaptive development (Gillespie-Lynch et al., 2012).

Individual differences in RJA are more closely associated with cognitive or verbal skills than chronological age for children on the spectrum. Indeed, reduced RJA may not be observed among children with autism who have attained a somewhat undefined developmental level (Leekam, Hunnisett, & Moore, 1998; Mundy, Sigman, & Kasari, 1994). This circumscribed developmental period within which RJA deficits are apparent in autism contrasts sharply with the difficulty even high functioning people on the spectrum have initiating joint attention (IJA: see Mundy & Newell, 2007 for a discussion of this).
Contradictory findings from participants of varied developmental levels and across diverse paradigms have led to a call for a developmental approach to the origins of RJA difficulties in autism (e.g. Nation & Perry, 2008). The first study in this dissertation highlights the long-term implications of the developmental window in which RJA impairments are evident in autism by examining associations between individual differences in early childhood RJA and changes in the skills and symptoms of individuals on the spectrum into adulthood. The second study focuses much earlier in development to see if an eye-tracking measure of RJA distinguishes between the infant siblings of children with autism (infant siblings), who are at heightened risk for developing autism (Bolton et al., 1994; Ozonoff et al., 2011; Ritvo et al., 1988), who do and do not develop autism, as well as a low-risk control group. Having explored longitudinal associations between RJA and subsequent skills and symptoms across development, I wished to examine the mechanisms underlying RJA deficits in autism. The last study in this series was designed to distinguish between three potential theoretical explanations for reduced RJA in autism: low-level difficulties with the type of gaze following that typically emerges within days of birth (Johnson et al., 2005; but see Chawarska et al., 2010), impaired integration of gaze cues and affect (Akechi et al., 2009; Kasari, Sigman, Mundy, & Yirmiya, 1990) and decreased recognition of the referential significance of gaze cues (Baron-Cohen, Baldwin, & Crowson, 1997; Gliga et al., 2012).

Because it is often associated with changes in cognitive and linguistic skills, increased understanding of how RJA develops among people with autism may provide guidance to interventionists. Interventions which ameliorate RJA (and often IJA) difficulties among children with autism yield improvements in language and social responsiveness (Jones, Carr, & Feeley, 2006; Kasari, Paparella, Freeman, & Jahromi, 2008; Whalen, Schreibman, & Ingersoll, 2006).
My hope is that this series of studies will help in the development of better screening measures for autism risk and will aid interventionists in deciding which key aspects of RJA might be the best targets of interventions for a given individual. I also hope that this series of studies brings forth the richness and complexity of autism. Autism is associated with both strengths to celebrate and difficulties to alleviate (see Kapp, Gillespie-Lynch, Sherman & Hutman, in press for a discussion of this). This series of studies focuses on an area of difficulty in autism. However, the same types of combinations of longitudinal studies across different stages of development combined with cross-sectional studies to better understand the mechanisms underlying key aspects of autism could be used to explore both the strengths and weaknesses associated with autism.

For future researchers, it may often be more efficient to move in the reverse direction to the one in which I moved in this dissertation. Validate measures one wishes to use to assess autism risk in infancy by starting first with a study to identify key differences between individuals who are and are not on the spectrum. But Roethke said it best: “I learn by going where I need to go. We think by feeling. What is there to know?”
Study 1: Early Childhood Predictors of the Social Competence of Adults with Autism

Abstract

Longitudinal research into adult outcomes in autism remains limited. Unlike previous longitudinal examinations of adult outcomes in autism, the twenty participants in the current study were evaluated across multiple assessments between early childhood ($M = 3.9$ years) and adulthood ($M = 26.6$ years). In early childhood, responsiveness to joint attention (RJA), language, and intelligence were assessed. In adulthood, the parents of participants responded to interviews assessing the adaptive functioning, autistic symptomatology and global functioning of their children. RJA and early childhood language predicted a composite measure of adult social functioning and independence. Early childhood language skills and intelligence predicted adult adaptive behaviors. RJA predicted adult non-verbal communication, social skills and symptoms. Adaptive behaviors changed with development, but symptoms of autism did not. Additional factors associated with adult outcomes are assessed.
Do early childhood intelligence, language, and joint attention predict the independence, adaptive abilities, and symptomatology of adults with autism? Although autism is a developmental disorder, few studies have tracked the same individuals across multiple stages of development into adulthood and none have assessed relationships between childhood joint attention and adult outcomes, or how well individuals can function independently. Joint attention, or the ability to align one’s own attention with the attention of another, is important to study as it is foundational for symbolic reference and is commonly impaired in autism (Mundy, Sullivan, & Mastergeorge, 2009). The current study is the first to assess whether early childhood joint attention skills predict adult social functioning in autism.

**Does RJA Predict Adult Outcomes?**

Participants in the current study were assessed at four time points from early childhood to adulthood. Findings from the first three time points have been reported in previous publications (McGovern & Sigman, 2005; Sigman & McGovern, 2005; Sigman & Ruskin, 1999; Siller & Sigman, 2002). In early childhood, both responsiveness to joint attention (RJA) and initiation of joint attention (IJA) were assessed. Both RJA and IJA are relevant variables to consider as RJA may index involuntary social orienting while IJA may require more intentional control (Mundy et al., 2007). While both RJA and IJA predicted expressive language gains one year after the first assessment, only RJA predicted intelligence quotient (IQ) gains from the first to the second assessment and receptive language at the third assessment (Sigman & McGovern, 2005; Sigman & Ruskin, 1999). Thus, in the current study we hypothesized that RJA would be associated with adult outcomes.

**Previous Longitudinal Research on Adult Outcomes in Autism**
While the majority of adolescents and adults with autism achieve limited independence and social relatedness (see Table 1: Billstedt, Gillberg, & Gillberg, 2005; Cederlund et al., 2008; Eisenberg, 1956; Gillberg & Steffenburg, 1987; Howlin, Goode, Hutton, & Rutter, 2004; Kanner, 1971; Larsen & Mouridsen, 1997; Lotter, 1974; Rutter, Greenfield & Lockyer, 1957), exceptions to this pattern have been reported (Eaves & Ho, 2008; Kobayashi, Murata, & Yoshinaga, 1982) particularly for individuals with autism who have higher IQs (Farley et al., 2009; Kanner, Rodriguez, & Ashenden, 1972) or individuals with Asperger syndrome (Cederlund et al., 2008; Engström et al., 2004; Larsen & Mouridsen, 1997). However, better outcomes for those with Asperger syndrome relative to those with autism may not be apparent when both groups have comparable IQs (Howlin, 2003).

Table 1  Previous Longitudinal Studies of Adult Social Functioning in Autism

<table>
<thead>
<tr>
<th>Authors</th>
<th>Final N</th>
<th>Source</th>
<th>Diagnostic Criteria</th>
<th>Mean Age (Range) in Years</th>
<th>% Male</th>
<th>Mean Initial IQ</th>
<th>Outcome Criteria</th>
<th>Overall Social Functioning</th>
<th>Childhood Predictor of Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eisenberg et al., 1956</td>
<td>53</td>
<td>Hospital: Before Behaviors</td>
<td>Intake: 6</td>
<td>FU: 15 (9-25)</td>
<td>79 NA</td>
<td>Vague</td>
<td>5% V/G, 22%</td>
<td>F, 73% V/P</td>
<td>Speech by 5</td>
</tr>
<tr>
<td>Lotter, 1974</td>
<td>32</td>
<td>Population screening: 1964</td>
<td>Creak</td>
<td>Intake: (8-10)</td>
<td>NA</td>
<td>67% &lt;50</td>
<td>Rutter &amp; Lockyer</td>
<td>13% V/G, 23%</td>
<td>Speech, IQ</td>
</tr>
<tr>
<td>Howlin et al., 2004</td>
<td>68</td>
<td>Hospital: 1950-1979 Rutter → DSM-IV</td>
<td>Intake: 7(3-16)</td>
<td>FU: 29 (21-49)</td>
<td>90 80</td>
<td>Howlin</td>
<td>22% V/G, 19%</td>
<td>PIQ, Speech by 5</td>
<td></td>
</tr>
<tr>
<td>Howlin et al., 2000</td>
<td>19</td>
<td>Hospital schools: Before 1971</td>
<td>Rutter</td>
<td>Intake: 7</td>
<td>100 NA</td>
<td>Mawhood</td>
<td>16% V/G, 10%</td>
<td>Vocabulary, PIQ</td>
<td></td>
</tr>
<tr>
<td>Gillberg &amp; Steffenburg, 1987</td>
<td>23</td>
<td>Population screening: Born</td>
<td>DSM-III</td>
<td>Intake: NA</td>
<td>74</td>
<td>20% &gt;70</td>
<td>Lotter</td>
<td>59% V/P</td>
<td>IQ &gt;50, Speech by 6</td>
</tr>
<tr>
<td>Billstedt et al., 2005</td>
<td>78</td>
<td>3 population screenings: Born</td>
<td>DSM-III</td>
<td>Intake: &lt; 10</td>
<td>70</td>
<td>26% &gt;70</td>
<td>Lotter</td>
<td>8% F, 16% R, IQ, Speech</td>
<td>75% V/P,</td>
</tr>
<tr>
<td>Authors</td>
<td>Final N</td>
<td>Source</td>
<td>Diagnostic Criteria</td>
<td>Mean Age (Range) in Years</td>
<td>% Male</td>
<td>Mean Initial IQ</td>
<td>Outcome Criteria</td>
<td>Overall Social Functioning</td>
<td>Childhood Predictor of Outcome</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------</td>
<td>---------------------------------</td>
<td>---------------------</td>
<td>---------------------------</td>
<td>--------</td>
<td>----------------</td>
<td>-------------------</td>
<td>--------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td></td>
<td>70 AS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AS: 22 (16-34),&lt;70</td>
<td>Aut: 83%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25 (16-36)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eaves &amp; Ho, 2008</td>
<td>48 Aut.</td>
<td>Clinic: Recruited</td>
<td>Rutter and DSM-I</td>
<td>Intake: 6 (3-12)</td>
<td>75</td>
<td>61% &gt; 50</td>
<td>Eaves &amp; Ho 21% V/G,</td>
<td>F, 46% V/P,</td>
<td>particularly VIQ</td>
</tr>
<tr>
<td></td>
<td>40 ASD</td>
<td>1974-1984,</td>
<td></td>
<td>FU: 11 (8-17)</td>
<td></td>
<td>AS: 78</td>
<td>Lotter</td>
<td>V/P</td>
<td>IQ,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AS: All &gt; 70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Aut: 33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Aut: 4 &lt; 70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9 AS</td>
<td>1949-1951</td>
<td></td>
<td></td>
<td></td>
<td>AS: 78</td>
<td></td>
<td>V/P</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Aut: 37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FU: AS: 39 (33-34), Aut: 37 (32-39)</td>
<td>62</td>
<td>22.1% &gt; 70</td>
<td>Kobayashi</td>
<td>F, 46% V/P,</td>
<td>IQ (all), Speech (males)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farley et al., 2009</td>
<td>41</td>
<td>Population screening: Born</td>
<td>Rutter and DSM-I</td>
<td>Intake: 7 (3-26)</td>
<td>93</td>
<td>94</td>
<td>Lotter</td>
<td>48% V/G, 14%</td>
<td>IQ, Change in IQ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14% V/G, 25%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>61% V/P</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FU: 16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kobayashi et al., 1992</td>
<td>197</td>
<td>Therapeutic agencies: Born</td>
<td>Rutter and DSM-I</td>
<td>Intake: 6 (SD=2.8)</td>
<td>84</td>
<td>22.1% &gt; 70</td>
<td>Kobayashi</td>
<td>37% V/G, 27%</td>
<td>IQ (all), Speech (males)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>before 1972.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FU: 22 (18-33)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Engström et al., 2003</td>
<td>32 AS</td>
<td>Population Health Records 1998</td>
<td>Rutter</td>
<td>Intake: 5</td>
<td>57</td>
<td>NA</td>
<td>Lotter</td>
<td>12% V/G, 75%</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>10 HFA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F, 12% V/P</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ruble &amp; Dalrymple, 1996</td>
<td>46</td>
<td>Dev. Dis. Center: Before 1996</td>
<td>Rutter</td>
<td>Intake: 5</td>
<td>72</td>
<td>26% &gt; 70</td>
<td>Vague</td>
<td>100% V/P</td>
<td>NA</td>
</tr>
</tbody>
</table>

a FU = Follow-Up

b V/G = Very/Good, F = Fair, R = Restricted but Acceptable, P/V = Poor/Very Poor
Types of Adult Outcomes

Identifying factors predictive of outcome across longitudinal studies of autism is complicated by variation in diagnostic criteria, IQ, age at initial and follow-up assessments, available early identification and intervention services, and the use of different and often subjective outcome variables (see Table 1: Kobayashi et al., 1982; Lord & Venter, 1992; Lotter, 1974; Lotter, 1978; Rutter & Lockyer, 1967; Venter, Lord, & Schopler, 1992). Three common measures of outcome are categorical assessments of independence and social relatedness (social functioning), adaptive behavior skills, and autistic symptoms.

While these outcome measures are often assessed indirectly via caregiver report, each provides unique insights into how individuals with autism develop into adulthood. Social functioning is a global measure of whether adults with autism are employed, have friends, and live independently. Despite variability in means of assessment across studies, social functioning is the most commonly used outcome measure across longitudinal studies of adult outcomes in autism and thus facilitates comparisons across studies. Longitudinal assessment of the adaptive behaviors of individuals with autism may allow educational and vocational opportunities to be tailored to individual needs (Carter et al., 1998; Freeman, Rahbar, Ritvo, Bice, Yokota & Ritvo, 1991; Volkmar, Sparrow, Goudreau, Cicchetti, Paul & Cohen, 1987). Developmental change in the symptoms of autism may provide insights into the natural course of the disorder, as well as supporting service planning (Fecteau, Mottron, Berthiaume, & Burack, 2003; Piven, Harper, Palmer, & Arndt, 1996; Seltzer, Krauss, Shattuck, Ormond, Swe & Lord, 2003).

Predictors of Adult Social Functioning
Social functioning in adults with autism has been related to speech before age 6 (Eisenberg, 1956; Kanner et al., 1972), early childhood IQ (Eaves & Ho, 2008; Farley et al., 2009), or a combination of language skills and IQ (Billstedt et al., 2005; Gillberg & Steffenburg, 1987; Howlin, Mawhood, & Rutter, 2000; Kobayashi et al., 1992; Lotter, 1974; Rutter et al., 1957). While non-verbal IQ (NVIQ) and language ability are only moderately related for individuals who are not intellectually disabled, around 70% of individuals with autism may be at least somewhat cognitively impaired and severe intellectual disability almost always co-occurs with impaired language (Rutter, 1970). Despite being related, speech and IQ may explain unique portions of the variance in outcome (Rutter et al., 1957).

**Predictors of Adaptive Behaviors and Symptoms**

Language abilities (Venter et al., 1992), IQ (Sigman & McGovern, 2005), and a combination of the two (Anderson et al., 2009; Szatmari, Bryson, Duku, Vaccarella, Zwaigenbaum, Bennett & Boyle, 2009) also predict adaptive behavior skills. While adaptive behaviors and IQ are often correlated in individuals with autism, adaptive behaviors are often lower than would be expected based upon IQ (Freeman et al., 1991; Venter, Lord & Schopler, 1992), particularly for non-intellectually disabled individuals (Bölte & Poustka, 2002). IQ may (Sigman & McGovern, 2005) or may not (Fecteau et al., 2003) predict improvement in symptoms with age.

Relationships between adaptive behaviors and symptoms may vary with IQ. While the social skills and social symptoms of more intellectually disabled children and adolescents are moderately correlated (Anderson, Oti, Lord, & Welch, 2009), they are less consistently related for higher-functioning individuals (Klin, Saulnier, Sparrow, Cicchetti, Volkmar, & Lord, 2006).
Participants in the current study had a mean early childhood IQ of 55; therefore, we expected social skills and social symptoms to be correlated across development.

In one of the few studies to assess adaptive behaviors and symptoms at multiple time points, Szatmari et al. (2009) used hierarchical linear modeling to delineate the trajectories of Vineland Adaptive Behavior Scales (VABS; Sparrow et al., 1984) scores and symptoms on the Autism Behavior Checklist (Krug, Arick, & Almond, 1980) of high functioning individuals with autism across four assessments from early childhood through adolescence. After verifying that participants were on the autism spectrum, a classification of autism or Asperger syndrome (AS) was conferred based on grammatical impairments between 6 and 8 years of age. While children with AS had better VABS scores across all domains and time points, growth in adaptive behaviors was independent of diagnosis and flattened out in late adolescence. NVIQ assessed at 5.5 years was related to VABS daily living skills and socialization but not communication scores. VABS daily living and socialization skills also improved for participants in the second and third assessments of our study, particularly those with higher IQs (Sigman & McGovern, 2005).

**Is Change in Language and Intelligence Predictive of Adult Outcomes?**

Do early childhood IQ and language impact adult outcomes simply by remaining consistent across development? Early childhood IQ is predictive of IQ in adolescence and adulthood in autism (Farley et al., 2009; Freeman et al., 1991; Sigman & McGovern, 2005). However, lack of change in group-level IQs can obscure substantial variation in the IQ of individuals (Cederlund et al., 2008; Farley et al., 2009; Lockyer & Rutter, 1969). While IQ may be less stable for individuals who never develop language (Lord & Schopler, 1991a; Rutter,
higher IQ may differentiate between initially non-verbal children who do and do not
develop language after age five (Rutter et al., 1967). Change in IQ from first assessment to
follow-up may predict better social functioning approximately 25 years after first assessment
(Farley et al., 2009). Language has often been assessed as either present or absent; however,
change in a continuous language measure may also be a powerful predictor of social functioning
outcomes.

**Does RJA Predict Change?**

While Szatmari et al. (2009) found that individuals with AS had fewer autistic symptoms
than those with autism across all assessments, symptoms decreased for both diagnostic groups
with age. Similarly, retrospective comparisons of the current and lifetime symptoms of
adolescents and adults on the ADI-R suggest that, for primarily low functioning populations as
well as for large samples of individuals with unspecified IQ, all ADI-R symptom domains (e.g.,
social, verbal and non-verbal communication, and repetitive behaviors) improve with age
(McGovern & Sigman, 2005; Seltzer et al., 2003b), while for higher functioning populations,
social and communicative symptoms (as quantified by the ADI-R) may decrease more than
repetitive behaviors (Fecteau et al., 2003; Piven et al., 1996). However, prospective comparisons
suggest that non-verbal communication may not improve with age (McGovern & Sigman, 2005;
Shattuck et al., 2007).

Szatmari et al. (2009) suggested that the absence of grammatical impairment, increased
VABS scores, and decreased autistic symptoms might all arise from a common developmental
precursor such as joint attention. Short-term longitudinal studies of children with autism indicate
that more frequent IJA, as indexed by gaze alternation, predicts reduced social and
communicative symptoms (Charman, 2003) and both IJA and RJA predict better expressive language (Kasari, Paparella, Freeman, & Jahromi, 2008; Sigman & Ruskin, 1999). Thus, IJA might be related to symptoms in adulthood while both types of joint attention may be related to linguistic competence, which in turn might influence adult adaptive abilities and independence.

However, RJA, but not IJA, during the first assessment was related to language at the third assessment in the current set of studies (Sigman & McGovern, 2005). Tantam (1992) postulated that the failure of a typically innate tendency to respond to joint attention (RJA) may be a central deficit in autism which makes individuals with autism more apt to focus on idiosyncratic rather than shared attention structures and less likely to learn word-object correspondences (Baldwin, 1991). From time point one to time point two of the current set of studies, 26% of the sample moved out of the intellectually disabled range, and those who did so exhibited more RJA during the first assessment than those who remained intellectually disabled (Sigman & Ruskin, 1999). Given its effects on language and cognitive development, we expected RJA to predict adult independence and adaptive skills, though we expected its effects to be reduced when changes in language and IQ were also included in analytic models.

Methods

Participants

The current report is based on interviews with the parents of twenty individuals with autism (M=26.6 years, SD = 3.8) who were assessed during three prior assessments when participants in the current report had a mean age of: 3.9 years (SD = 1.2 years), 11.7 years (SD = 3.2), and 18.3 years (SD = 3.6). See Table 2 for participant characteristics across time points. While the first three assessments included standardized testing, behavioral observations,
interviews and questionnaires, the current assessment consisted solely of tape-recorded parental interviews and questionnaires. As many participants had moved since the last assessment, interviews were conducted over the phone, although one parent elected to do the interviews in person.

Table 2  *Descriptive Statistics, N:Mean (SD), for Final Adult Sample across Time*

<table>
<thead>
<tr>
<th></th>
<th>Time Point 1</th>
<th>Time Point 2</th>
<th>Time Point 3</th>
<th>Time Point 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chronological Age</strong></td>
<td>20: 3.9 (1.2)</td>
<td>19: 11.7 (3.2)</td>
<td>20: 18.3 (3.6)</td>
<td>20: 26.6 (3.8)</td>
</tr>
<tr>
<td><strong>Mental Age</strong></td>
<td>20: 2.2 (1.2)</td>
<td>NA</td>
<td>20: 8.1 (6.4)</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Language Age</strong></td>
<td>20: 1.7 (0.9)</td>
<td>NA</td>
<td>20: 5.0 (4.4)</td>
<td>NA</td>
</tr>
<tr>
<td><strong>DQ</strong></td>
<td>20: 54.7 (15.5)</td>
<td>NA</td>
<td>20: 44.8 (34.6)</td>
<td>NA</td>
</tr>
<tr>
<td><strong>ESCS:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>IJA</strong></td>
<td>18: 8.79 (6.4)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>RJA</strong></td>
<td>18: .58b (.35)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>VABS Raw Scores:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Socialization</strong></td>
<td>NA</td>
<td>19: 59.5 (30.5)</td>
<td>17: 69.6 (37.9)</td>
<td>20: 67.3 (34.5)</td>
</tr>
<tr>
<td><strong>Communication</strong></td>
<td>NA</td>
<td>19: 69.6 (40.9)</td>
<td>17: 79.4 (46.0)</td>
<td>20: 80.0 (45.5)</td>
</tr>
<tr>
<td><strong>Daily Living Skills</strong></td>
<td>NA</td>
<td>19: 89.8 (36.9)</td>
<td>17: 111.5 (43.2)</td>
<td>20: 122.8 (43.6)</td>
</tr>
<tr>
<td><strong>ADI-R Algorithm:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Social</strong></td>
<td>NA</td>
<td>19: 16.4 (7.5)</td>
<td>16: 11.1 (6.8)</td>
<td>19: 13.3 (5.8)</td>
</tr>
<tr>
<td><strong>Non-verbal Communication</strong></td>
<td>NA</td>
<td>19: 4.3 (3.1)</td>
<td>16: 3.6 (3.3)</td>
<td>19: 4.0 (3.2)</td>
</tr>
<tr>
<td><strong>Restricted and Repetitive</strong></td>
<td>NA</td>
<td>19: 5.2 (2.0)</td>
<td>16: 3.6 (2.1)</td>
<td>19: 4.0 (2.4)</td>
</tr>
</tbody>
</table>

*Frequency of initiations of joint attention over entire administration of ESCS*

*bProportion of responses to joint attention relative to presses for joint attention*

Seventy children with autism were first diagnosed in the late 1970s and early 1980s according to DSM-III criteria (Sigman & Ruskin, 1999). Fifty one (73% of the original sample) participated in the second assessment when the ADI was used to verify diagnosis for all current participants except one who missed the cut-off for repetitive behaviors by 1 point and one who
did not participate in the second assessment (Sigman & Ruskin, 1999). Of those two individuals, one met criteria for autism and one met criteria for PDD-NOS on the ADOS during the third assessment. Forty eight (68% of the original sample) participated in the third assessment (McGovern & Sigman, 2005). Relationships between maternal behaviors and the development of a subset of the participants in the current study across the first three assessments were also reported by Siller and Sigman (2002). Twenty participated in the current follow-up study (29% of the original sample). Independent samples t-tests revealed that participants in the current assessment did not differ from the 50 participants in the first assessment who were lost to follow-up in terms of chronological age (p=.90) or mental age (p=.21) at first assessment. However, participants assessed during the current assessment had significantly higher developmental quotients (Current Participants: M= 54.65; Lost to Follow-Up: M=47.18; p=.032) and marginally higher language abilities (Current Participants: M= 20.13; Lost to Follow-Up: M=14.70; p=.054) at first assessment than participants who did not return for this follow-up assessment.

Twenty-eight participants from the third assessment did not participate in the current study for the following reasons: parents of 3 participants had died, 1 participant had died, 1 parent declined to participate, 20 participants could not be located, and 3 returned the consent form but did not respond to calls. While the previous assessment included 6 females and 42 males, the current sample was composed entirely of males: 13 Caucasian, 4 African American, 2 Asian, and 1 Hispanic.

Two participants did not complete the ESCS during the first assessment. The VABS was not completed during the second assessment for one participant and during the third for three participants. Six participants did not complete language and intelligence testing during the second assessment; therefore, the effects of changes in language and mental age on adult
outcome measures were assessed by relating scores at time point one to scores at time point three. The ADI-R was not completed for three participants during the third assessment and for one participant during the current assessment. The latter participant reported that her work schedule was too hectic to complete the ADI-R.

**Early Social Communication Scales (ESCS)**

Administered during the first assessment, the Early Social Communication Scales (Mundy, Hogan, & Doehring, 1996) is a structured observation of non-verbal communicative abilities including IJA (the frequency with which a child uses eye contact, pointing, and showing to initiate shared attention) and RJA (the proportion of prompts to elicit RJA when the child follows the experimenter’s gaze and pointing gestures).

**Intelligence Assessments**

Based on ability level, either the Cattell Scales of Development (15 participants) or the Stanford-Binet (Terman & Merrill, 1973) was given during the first assessment. At time 3, eleven participants were administered the Stanford-Binet 4th edition (Thorndike, Hagen, & Sattler, 1986) and nine received the Mullen Scales of Early Learning (Mullen, 1995). All tests yielded mental age equivalents (MA) which were divided by chronological age to yield developmental quotients (DQ).

**Language Assessments**

Language was measured in early childhood with the Reynell Scales of Language Ability (Reynell, 1977). During the third assessment, participants with limited speech were administered the Reynell, while the Clinical Evaluation of Language Fundamentals-Revised (CELF-R; Semel, Wiig, & Secord, 1987) was used for those with fluent speech. Each test yielded a language age equivalent (LA).
**Autism Diagnostic Interview-Revised (ADI-R)**

The ADI-R is a standardized, semi-structured caregiver interview that provides a diagnostic algorithm for the DSM-IV (American Psychiatric Association, 1993) definition of autism (Lord, Rutter, & Le Couteur, 1994). The ADI-R was administered by clinicians who had established second-degree reliability with the UMACC ADI-R training site on training videotapes. ADI-R questions assess the social, verbal, and non-verbal communicative symptoms of autism as well as restricted and repetitive behaviors at both the time of interview (“current”) and between the ages of four and five (“ever”).

Because the ADI-R was administered during the second and third assessments, only current functioning was evaluated during the current (fourth) assessment. Algorithm items which do not yield current ratings were excluded from analysis (Boelte & Poustka, 2002; Fecteau et al., 2003; Howlin, 2000; Lord et al., 1997; McGovern & Sigman, 2005). Seven participants had insufficient speech to assess verbal symptoms, so the verbal domain was not analyzed. As specified in the diagnostic algorithm, all ratings of 3 were transformed into 2 for analysis (Lord et al., 1994). Higher scores indicate greater symptom severity.

**Vineland Adaptive Behavior Scales (VABS) Interview Edition**

The VABS is a semi-structured caregiver interview assessing self-sufficiency across three domains: communication, socialization, and daily living skills (Sparrow et al., 1984), which was administered during assessments two, three, and four. Domain raw scores can be converted into standard scores or age equivalents. Because age equivalence scores may be misleading due to lack of comparability in range across domains and standard scores may be inappropriate for individuals with autism (Carter et al., 1998), analyses were performed on raw scores.
Overall Social Functioning

For the current study, we used a composite rating on a 5-point scale of overall social functioning based on employment, living situation, and friendships (from Howlin et al., 2004). Parents were asked a set of questions about their child’s level of functioning (see Appendix for questions asked and coding scheme). This composite rating was chosen as it is similar to outcome measures used in many other studies (See Table 1: Cederlund et al., 2008; Engström et al., 2004; Gillberg & Stefenburg, 1987; Larsen & Mouridsen, 1997; Lotter, 1978; Kobayashi et al., 1992; Billstedt et al. 2005, Rutter et al., 1967). According to this composite, the criteria for a “Very Good” rating included residential and employment independence as well as some friendships. A “Good” outcome signified either paid or voluntary employment with some degree of support in daily living and some friendships or acquaintances. Individuals rated as having a “Fair” outcome had achieved some supported independence and had acquaintances but no close friendships. A “Poor” outcome referred to individuals who required a high level of support and had few social contacts. A “Very Poor” outcome was given if the individual was living in a hospital. The first and second author discussed and reached agreement on all social functioning scores. A research assistant independently coded parent responses and attained 100% agreement with the social functioning scores assigned by the authors.

Results

Due to the small sample size, we regard the following analyses as exploratory and report partial correlation values as well as significance levels. DQ and LA were analyzed separately because separate estimates of DQ based on verbal and non-verbal mental age were not available. Indeed, DQ and LA at first assessment were highly correlated, $r (18) = .68$, $p = .001$. 

17
Chronological age at first assessment was entered into, and not significant, in all regressions except those which included DQ. CA was not entered into analyses which included DQ because DQ is defined by dividing MA by CA. Hierarchical linear modeling and regressions revealed similar predictive relationships between skills at first assessment and adult abilities and symptoms, so regressions are reported for ease of comprehension. Adaptive social skills as assessed by the VABS and social symptoms as assessed by the ADI-R were concurrently correlated \( (p=.001) \) at assessments two \( r(17)= -.71 \), three \( r(14)=-.80 \) and four \( r(17)= -.84 \).

Early childhood variables examined in relation to each of the outcome variables included developmental quotient at time 1 (DQ1), language age at time 1 (LA1), and early childhood RJA and IJA. Change in both LA and DQ from time 1 to time 3 was assessed in relation to each of the outcome variables by adding language (LA3) and developmental quotient (DQ3) from time 3 to models containing time LA1 or DQ1. A difference score indicating change in LA or DQ from time 1 to time 3 was also calculated to examine relationships between changes in abilities and a categorical outcome variable, adult social functioning. To test for possible mediators between RJA and outcome variables, relationships between RJA and change scores were examined: RJA predicted change in language skills from time 1 to time 3 \( (\beta = .790, t(16) = 3.696, pr=.690, p =.002) \) and change in DQ from time 1 to time 3 \( (\beta = .685, t(16) = 3.242, pr=.642, p =.005) \).

Details of the regression analyses described below are summarized in Table 3.
**Table 3 Beta Values of Simple Regressions Relating Early Childhood Predictors to Adult Outcomes After Controlling for Chronological Age**

<table>
<thead>
<tr>
<th>Childhood predictor →</th>
<th>IJA T1</th>
<th>RJA T1</th>
<th>DQ T1</th>
<th>LA T1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VABS:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Socialization</td>
<td>.256</td>
<td>.538*</td>
<td>.509*</td>
<td>.857**</td>
</tr>
<tr>
<td>Communication</td>
<td>.172</td>
<td>.501</td>
<td>.522*</td>
<td>.914 ***</td>
</tr>
<tr>
<td>Daily Living Skills</td>
<td>.276</td>
<td>.482</td>
<td>.480*</td>
<td>.824**</td>
</tr>
<tr>
<td><strong>ADI-R:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social</td>
<td>-.417</td>
<td>-.603*</td>
<td>-.379</td>
<td>-.721*</td>
</tr>
<tr>
<td>Non-verbal Communication</td>
<td>-.431</td>
<td>-.797**</td>
<td>-.315</td>
<td>-.569</td>
</tr>
<tr>
<td>Restricted/Repetitive Behaviors</td>
<td>-.302</td>
<td>-.378</td>
<td>-.246</td>
<td>-.327</td>
</tr>
</tbody>
</table>

*α< .05
*α< .01
**α< .001

**Social Functioning**

The percentage of participants classified into each level of social functioning was as follows: “Very Good” = 20%, “Good” = 10%, “Fair” = 20%, and “Poor” = 50% (see Table 4). Because social functioning is a categorical outcome measure, Spearman correlations were used to examine relationships between predictors and social functioning. Social functioning was related to LA1 ($\rho(18) = -.843, p < .001$), RJA ($\rho(16) = -.798, p < .001$), LA3-LA1 ($\rho(18) = -.866, p < .001$) and DQ3-DQ1 ($\rho(18) = -.825, p < .001$). However, social functioning was unrelated to DQ1 ($p = .080$) or IJA ($p = .125$). Thus, both early childhood language and RJA predicted adult social functioning. While change in language and DQ were also predictive of adult social functioning, a direct test to determine if change in skills mediated the relationship between RJA and social functioning was not conducted because of small sample size and because social functioning is a categorical variable.
<table>
<thead>
<tr>
<th>Participant</th>
<th>Independence</th>
<th>Type of Work</th>
<th>Friendship</th>
<th>Overall Functioning</th>
<th>Seizures</th>
<th>Medications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family Home: Can go out alone</td>
<td>Full-time maintenance work at parents' day care</td>
<td>Close friend, shares common interests</td>
<td>Very good</td>
<td>No</td>
<td>For attention</td>
<td></td>
</tr>
<tr>
<td>Family home: Manages own budget</td>
<td>Full-time medical filing clerk</td>
<td>Multiple friends and has dated</td>
<td>Very good</td>
<td>No</td>
<td>For anxiety</td>
<td></td>
</tr>
<tr>
<td>Own apartment in different state than parents</td>
<td>Full-time manager of small airline</td>
<td>Multiple friends, no dating</td>
<td>Very good</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Family Home: Can go out alone</td>
<td>Full-time work for Coca Cola and just earned AA</td>
<td>Has friends but they introduced him to a gang and took advantage of him</td>
<td>Very good</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Family home: Supervised in community</td>
<td>Community college: Studying to be history teacher</td>
<td>Extends interest-based friendships outside group situations</td>
<td>Good</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Family home: Looking for apartment</td>
<td>In college: Studying the environmental effects of the workspace</td>
<td>Acquaintances in group situations</td>
<td>Good</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Own apartment: Weekend staff</td>
<td>Part-time supported employment: Art production</td>
<td>No friends</td>
<td>Fair</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Family home: Always supervised</td>
<td>Sheltered employment at Community Service Center: Money changing</td>
<td>No friends</td>
<td>Fair</td>
<td>Yes</td>
<td>For blood pressure, cholesterol, stomach pain, epilepsy</td>
<td></td>
</tr>
<tr>
<td>Family home: Supervised in community</td>
<td>Custodial work at program: Cleaning pews and shredding paper</td>
<td>No friends</td>
<td>Fair</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Own apartment: Help with cleaning and taxes</td>
<td>Not employed</td>
<td>No friends</td>
<td>Fair</td>
<td>No</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Group home: Can go out alone</td>
<td>Sheltered workshop part-time</td>
<td>No friends</td>
<td>Poor</td>
<td>No</td>
<td>Mood stabilizer</td>
<td></td>
</tr>
<tr>
<td>Participant</td>
<td>Independence</td>
<td>Type of Work</td>
<td>Friendship</td>
<td>Overall Functioning</td>
<td>Seizures</td>
<td>Medications</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------</td>
<td>--------------</td>
<td>------------</td>
<td>---------------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>12</td>
<td>Group home: Can go out alone</td>
<td>Sheltered workshop part-time</td>
<td>No friends</td>
<td>Poor</td>
<td>No</td>
<td>Mood stabilizer</td>
</tr>
<tr>
<td>13</td>
<td>Group home: Always supervised</td>
<td>Not employed</td>
<td>No friends</td>
<td>Poor</td>
<td>Yes</td>
<td>For behaviors, epilepsy</td>
</tr>
<tr>
<td>14</td>
<td>Family home: Mom and caregiver supervise</td>
<td>Not employed</td>
<td>No friends</td>
<td>Poor</td>
<td>No</td>
<td>For behaviors, anxiety, depression</td>
</tr>
<tr>
<td>15</td>
<td>Family home: Weekend caregiver and family supervision</td>
<td>Not employed</td>
<td>No friends</td>
<td>Poor</td>
<td>No</td>
<td>Antipsychotics</td>
</tr>
<tr>
<td>16</td>
<td>Group home: Always supervised</td>
<td>Sheltered employment: Sorting things and loading water bottles</td>
<td>No friends</td>
<td>Poor</td>
<td>No</td>
<td>Multiple antipsychotics, mood stabilizers, anxiolytics</td>
</tr>
<tr>
<td>17</td>
<td>Group home: Always supervised</td>
<td>Not employed</td>
<td>No friends</td>
<td>Poor</td>
<td>No</td>
<td>For aggression, mood, Tourrettes, insomnia</td>
</tr>
<tr>
<td>18</td>
<td>Group home: Can go out alone</td>
<td>Not employed</td>
<td>No friends</td>
<td>Poor</td>
<td>No</td>
<td>Mood stabilizers</td>
</tr>
<tr>
<td>19</td>
<td>Family home: Supervised in community</td>
<td>Not employed</td>
<td>No friends</td>
<td>Poor</td>
<td>Yes</td>
<td>For epilepsy</td>
</tr>
<tr>
<td>20</td>
<td>Group home: Constant supervision</td>
<td>Supported program: food preparation, filing, and paper shredding</td>
<td>No friends</td>
<td>Poor</td>
<td>No</td>
<td>Mood stabilizer</td>
</tr>
</tbody>
</table>

**VABS Scores**

Daily living skills improved \( F (2, 30) = 15.442 <.001 \) overall and from time three to four \( t(16)=4.986, p<.001 \). When entered into simple regression models, LA1 accounted for 40% and DQ1 accounted for 19% of the variance in raw scores on the daily living skills domain.
at time four. RJA and IJA were unrelated to daily living skills. A model containing LA3 ($\beta = .609, t (16) = 3.434, pr = .651, p = .003$) and LA1 ($p = .127$) explained 63% of the variance in daily living skills. DQ3 ($\beta = .902, t (17) = 5.794, pr = .815, p < .001$) and DQ1 ($p = .654$) explained 71% of the variance in daily living skills. While early childhood LA and DQ (i.e., LA1 and DQ1) predicted adult daily living skills, change in LA and DQ between time 1 and time 3 were stronger predictors of daily living skills than baseline measures were.

Improvement in Communication skills ($F(2,30)=3.405, p = .047$) was significant across the second, third and fourth assessments. Follow-up t-tests indicated significant improvements in communication skills from time two to time four ($t (19) = -2.233, p = .039$), but not from time three to time four ($t (16) = -1.969, p = .067$). LA1 explained 49% and DQ1 explained 23% of the variance in raw scores on the communication domain at time four while RJA and IJA were not related to communication skills. A model containing LA1 ($\beta = .495, t (16) = 2.288, pr = .497, p = .036$) and LA3 ($\beta = .582, t (16) = 3.638, pr = .673, p = .002$) explained 70% of the variance in communication skills. DQ3 ($\beta = .887, t (17) = 6.002, pr = .824, p < .001$) and DQ1 ($p = .899$) predicted 74% of the variance in communication skills. Early childhood language (i.e., LA1) accounted for additional variance in adult communication skills not explained by change in language from the first to the third assessment. However, change in DQ from time 1 to time 3 appeared to mediate the relationship between early childhood DQ and adult communication skills.

Social skills did not differ between the second, third, and fourth assessments ($F(2,30) = .273, p = .763$). LA1 explained 51%, DQ1 explained 22%, and RJA explained 34% of the variance in raw scores on the socialization domain at time four, while IJA was not significantly related to social skills. When LA1 ($\beta = .687, t (14) = 2.446, pr = .547, p = .028$) and
early childhood RJA (p=.232) were simultaneously entered into a regression model, it accounted for 50% of the variance in social skills. A model containing LA3 (β = .588, t (13) = 2.522, pr=.573, p =.025), LA1 (β = .508, p =.062), and RJA (β = -.111, p =.662) explained 64% of the variance in social skills. When early childhood RJA (β = .513, t (15) = 2.394, pr=.526, p =.030) and DQ1 (p=.245) were simultaneously entered into a regression model, the model accounted for 38% of the variance in social skills. A model containing DQ3 (β = .798, t (14) = 3.414, pr=.674, p =.004), DQ1 (β = -.018, p =.923), and RJA (β = .067, p =.755) explained 63% of the variance in social skills. Thus, all early childhood variables except IJA were related to adult social skills. While relationships between RJA and adult social skills may have been mediated by LA1, associations between LA1 and adult social skills may in turn have been mediated by change in language skills from time one to time three. Additionally, change in DQ from time 1 to time 3 appeared to mediate relationships between RJA and adult social skills. Thus, RJA influenced adult social skills through concurrent relationships with early childhood language and predictive associations with change in DQ.

**ADI-R Symptoms**

Social interaction algorithm scores changed between the second, third, and fourth assessments (F (2, 28) =4.829, p=.016). T-tests indicate that symptoms decreased from time two to time three (t(14)=2.94, p=.011), and then increased from time three to time four (t(15)=-2.20, p=.044). Social symptoms did not differ between the second and fourth assessments. RJA explained 33% and LA1 explained 30% of the variance in social symptoms while DQ1 and IJA were unrelated to social symptoms. When early childhood RJA (p =.135) and LA1 (p=.057) were simultaneously entered into a regression model, the model accounted for 46% of the variance in social functioning. While a model containing LA1, LA3, and RJA explained 42% of
the variance in social symptoms, none of the predictors was significantly related to social symptoms. Thus, RJA and LA1 accounted for overlapping aspects of adult social symptoms and there was no evidence that change in language mediated the relationship between RJA and social symptoms.

Neither non-verbal communication algorithm scores ($F(2,28)=.408, p=.669$) nor restricted and repetitive behavior algorithm scores ($F(2,28)=2.789, p=.079$) changed across the second, third, and fourth assessments. RJA accounted for 47% of the variance in non-verbal symptoms at time four, while DQ1, LA1, and IJA were not related to non-verbal communication. No early childhood scores predicted restricted and repetitive behaviors.

**Discussion**

The social functioning outcomes of participants in the current study are comparable to those reported by Eaves and Ho (2008) for another population born in the 1970s and 1980s with similar intelligence levels. Both studies suggest that adult social functioning outcomes for individuals with autism may be improving gradually. Additionally, somewhat better outcomes were also noted when comparing longitudinal studies conducted after 1980 to those conducted prior to 1980 (Howlin and Goode, 1998). This trend is probably due to the increasing availability of services, particularly as similar outcomes were obtained for individuals born prior to 1972 who participated in intensive community based interventions (Kobayashi et al., 1992).

Selective attrition of particularly low functioning individuals with autism may have inflated the proportion of participants with better outcomes in the current study. While the average intelligence level at first assessment of the twenty participants in the current report was quite low ($M= 54.65$), it was higher than the average intelligence level of the fifty participants
who were lost to attrition (M=47.18). Although other studies documenting slight increases in positive outcomes have not lost as many participants to attrition as were lost in the current study, comparisons between participants who were and were not lost were not reported in those studies (Eaves & Ho, 2008; Kobayashi et al., 1992). Therefore, as Eaves and Ho also acknowledged, increasingly positive outcomes in more recent longitudinal studies of adult outcomes in autism may be at least partially due to selective attrition of lower functioning participants.

Language skills and RJA, but not intellectual functioning, predicted adult social functioning. Intellectual functioning may have been less prognostic than in other longitudinal studies because the average age of first assessment was quite young in this study (see Table 2), NVIQ was not assessed, and/or intelligence may discriminate best among those with poor and very poor outcomes (Rutter et al., 1957). Moreover, very poor outcomes are no longer as prevalent due to improvement in services, as well as deinstitutionalization, or the ongoing migration of disabled populations from institutions to community residential arrangements. Some of the predictive potential of language ability (in terms of social functioning) appears to be due to its relationship with RJA, which may have scaffolded changes in DQ and LA. However, it was not possible to determine if predictive relationships between RJA and social functioning were mediated by change in skills with development.

While early childhood LA and DQ were related to all VABS domains, RJA was only related to the social skills domain. Indeed, relationships between RJA and social skills appeared to be mediated by change in intelligence from time one to time three. RJA was also related to social symptoms and non-verbal communication in adulthood. Thus, early childhood RJA may be particularly predictive of social behaviors in adulthood. The lack of a relationship between
IJA and any of the outcome measures may demonstrate the prognostic value of more involuntary non-verbal communicative behaviors (Mundy et al., 2007) for adult social outcomes.

Factors other than RJA contributed to relationships between changes in DQ and LA and adult outcomes, as evidenced by the finding that changes in DQ and LA from time one to time three predicted VABS daily living and communication skills in the absence of direct connections between RJA and these skills. Maternal behaviors, such as synchrony, were associated with increases in RJA, IJA, and language for many of the participants in the current study across the first, second and third assessments (Siller & Sigman, 2002). Thus, parental behaviors which were not assessed in the current analyses may have also influenced adult outcomes.

The robust relationships between changes in DQ and LA from a mean age of 4 to a mean age of 18 and both VABs scores and social functioning illustrate several important points. First, this finding highlights the importance of skills such as RJA that facilitate learning from others. Second, these results illustrate the potentially powerful impact of early interventions and parental behaviors which promote linguistic and cognitive growth (Kasari, Paparella, Freeman, & Jahromi, 2008; Rogers, 1996; Siller & Sigman, 2002). Finally, our findings suggest that clinicians should be cautious when counseling parents on what to expect in the future based on early childhood abilities. The latter point is buttressed by the finding that the most consistent predictors of adult outcomes in this study were not early childhood characteristics, but changes in language and mental age between the first and third assessments.

When using VABS raw scores rather than the age equivalents used by McGovern and Sigman (2005), only daily living skills show strong evidence of improvement across development. Arguably, Daily Living Skills is the VABS domain which is the most amenable to
explicit instruction. Increases in VABS socialization scores in younger populations than the one studied here may be due to greater availability of effective interventions for younger cohorts (Anderson et al., 2009). Possibly due to a small sample size and low power, limited evidence of change in ADI-R symptoms with development was evident in the current sample. Our results suggest that, even when symptoms and abilities are correlated, they may develop differently.

Several factors may limit the generalizability of these findings. The small sample size, reliance on telephone interviews, and biased gender ratios are common limitations across longitudinal studies (Eaves & Ho, 2008; Larsen & Mouridsen, 1997; Mawhood et al., 2000; Szatmari et al., 1989). Reliance on parent report of adult outcomes reduced the depth of information available and may have introduced recall biases particularly about those individuals who were no longer living with family. Direct assessment of the individuals with autism themselves may have allowed for more detailed comparisons between characteristics assessed in early childhood and again in adulthood. However, telephone interviews were selected for practical reasons, as Eaves and Ho (2008) also noted. For example, many participants had moved out of the state. Additionally, while participants from earlier stages of this study did not differ from current participants in terms of age at first assessment, they did differ in terms of DQ and LA in a manner suggestive of selective attrition of lower functioning individuals.

Environmental characteristics, such as socioeconomic status, available services, and parental behaviors, were not assessed and may be related to the outcomes of interest. Furthermore, the generalizability of these results to children who are newly diagnosed may be limited by changes in diagnostic criteria, a primarily low-IQ sample, and changes in the quality and quantity of available interventions. Additional individual characteristics which we did not assess, such as theory of mind and executive function, may also have influenced adult outcomes.
However, joint attention reflects emerging social cognition and may be a precursor to theory of mind (Charman, Baron-Cohen, Swettenham, Baird, Cox & Drew, 2000). Concurrent relationships between joint attention and executive function in early childhood suggest that difficulties recognizing stimulus-reward contingencies may influence the development of joint attention and executive functions in autism (Dawson et al., 2002). Many aspects of executive functioning are concurrently related to the adaptive behavior skills of children with autism (Gilotty, Kenworthy, Sirian, Black & Wagner, 2002) and thus might be expected to predict changes in adaptive behavior. Future research in this area should assess relationships between joint attention and executive function longitudinally, particularly in relation to adult outcomes.

While relationships between RJA, language functioning, and adult outcomes illustrate the importance of joint attention interventions, the outcome measures used in this study were based on caregiver perceptions of outcomes. Individuals classified as having a “poor” outcome may experience life as happy and valued members of their communities (Ruble & Dalyrmple, 1996). Future longitudinal studies of outcome in autism would benefit from multidimensional measures both during initial assessment and follow-up. Measures that we recommend for future studies include early childhood RJA, measures of executive function, detailed information about education and interventions, and multiple outcome measures, including direct interviews that allow the individuals with autism themselves to describe and evaluate their own social and adaptive functioning.
Appendix

Calculating Social Functioning

**Independence**: Interviewer asked: “Where/with whom does your child live?

0 = living independently

1 = semi-sheltered accommodation (or still at home) but with a high degree of autonomy

2 = living with parents, some limited autonomy

3 = in residential accommodation with some limited autonomy

4 = specialist autistic or other residential accommodation with little or no autonomy

5 = in hospital care or at home because nowhere else would accept the individual.

**Work**: Interviewer asked the following set of questions from highest to lowest level of employment until one was endorsed.

0 = Is your child employed or self employed?

1 = Is your child involved in voluntary work or job training?

2 = Is your child involved in supported or sheltered employment?

3 = Is your child in a special center or not employed?

**Friendship**: This was calculated from parent response to question 65 of the ADI-3. Interviewer asked: Does your child have any particular friends or a best friend?

0 = One or more friendships defined by mutual reciprocity/responsiveness.

1 = One or more relationships outside of prearranged situations but limited in terms of restricted interests or reciprocity.

2 = Relationships involving seeking contact but only in group situations.

3 = No peer relationships involving selectivity or sharing.
**Overall Social Functioning:** Assigned based on summed composite of scores on the above three domains.

0 = Very Good outcome- i.e. achieving a high level of independence, having some friends or a job (total from all 3 areas above 0-2)

1 = Good outcome- generally in work but requiring some degree of support in daily living; some friends and acquaintances (total 3-4)

2 = Fair outcome- has some degree of independence, and although requires support and supervision does not need specialist residential provision; no close friends but some acquaintances (total 5-7)

3 = Poor outcome – requiring special residential provision/high level of support; no friends outside of residence (total 8-10)

4 = Very Poor- needing high-level hospital care, no friends; no autonomy (total 11)
Study 2: Eye-Tracking Gaze Following in Infants at Risk for Autism

Abstract

We evaluated whether an eye-tracking measure of social attention, specifically interest in dyadic cues, and gaze following could identify early markers of autism and the broader autism phenotype. Longitudinal assessments of infants at high- and low-risk for autism were conducted at 6, 12, 18 and 24 months of age. Infants who were later diagnosed with autism attended less to dyadic cues at 6 or 12 months than typically developing infants. Attention to dyadic cues increased more for toddlers with autism than typically developing children. No differences in gaze following were observed between infants with and without autism. High-risk infants attended more to the stimuli overall than low-risk infants. This difference could be attributable to the different contexts in which the two risk groups were tested. Findings suggest that attention to dyadic cues is a potential marker of autism in infancy. Future studies should conduct multiple assessments of social attention in infancy in order to elucidate whether aberrant or delayed trajectories of social attention are associated with autism in infancy.
Are reduced social attention and decreased gaze following risk markers of autism in infancy? While typically developing infants orient toward direct gaze and exhibit low level aspects of gaze following within moments of birth (Farroni, Massaccesi, Pividori & Johnson, 2004), children with autism attend to social information and respond to joint attention (defined behaviorally as gaze or point following) less than typically developing children and children with other disabilities (Chawarska, Macari & Shic, in press; Dawson et al., 2004; Sigman & Ruskin, 1999). How and when symptoms of autism emerge is likely to vary greatly among individuals with autism (e.g. Bryson et al., 2007; Lord, Luyster, Guthrie, & Pickles, in press). Given this variability in developmental trajectories within autism and the benefits of early intervention (e.g. Dawson, 2008), sensitive measures capable of identifying emerging symptoms of autism are needed. The first aim of the current study was to determine if an eye-tracking measure of social attention and responsiveness to joint attention (RJA) differentiates infant siblings of children with autism, who are at heightened risk for developing autism (e.g. Ozonoff et al., 2011), who go on to develop autism from infants siblings who do not develop autism.

Atypical responses to social information may also be a characteristic of the broader autism phenotype, or subclinical characteristics of autism present among individuals with autism and some of their relatives (Dawson, Webb, Wijsman, Schellenberg, Estes, Munson, & Faja, 2005; Piven, Palmer, Jacobi, Childress, & Arndt, 1997). Greater understanding of the broader autism phenotype could elucidate relations between risk and resilience in autism, as well as helping to identify behavioral endophenotypes that might be more directly associated with specific genes than the suite of behaviors required to receive an autism diagnosis (Losh & Piven,
The second aim of this study was to assess whether reduced social responsiveness in infancy is a characteristic of the broader autism phenotype.

**Social Risk Markers of Autism in Infancy**

Both retrospective studies of children diagnosed with autism and prospective studies of infant siblings indicate that social and non-social atypicalities of attention, perception and motion within the first few years of life are often apparent in infants who develop autism (Baranek, 1999; Osterling & Dawson, 1994; Pierce, Conant, Hazin, Stoner & Desmond, 2011; Zwaigenbaum et al., 2005). Retrospective studies have documented reduced social attention (including attention to faces) by 6 months of age (Maestro et al., 2005), reduced RJA and social referencing between 6 and 10 months (Clifford & Dissanayake, 2008) and decreased attention to people and orienting to one’s name between 6 and 12 months (Baranek, 1999; Osterling & Dawson, 1994; Osterling et al., 2002; Werner et al., 2002). Interestingly, infants who exhibited reduced social attention at 6 months no longer did at 12 months (Maestro et al., 2005). The authors speculated that differences in social attention between infants with and without autism decrease as typically developing infants become more interested in objects but that early dyadic deficits might impact later triadic skills such as joint attention.

Unlike retrospective studies, prospective studies have documented few behavioral differences between infants who do and do not develop autism prior to 12 months of age (see Rogers, 2009 for a discussion of this). Prospective designs have revealed that reductions in the following social behaviors at 12 months of age are associated with subsequent autism diagnoses: social attention (Zwaigenbaum et al., 2005), including orienting to name (Nadig, Ozonoff, Young, Rozga, Sigman, & Rogers, 2007) and gaze to an examiner’s face (Ozonoff et al., 2010),
particularly during potentially affective interactions (Hutman et al., 2010; 2011). Interestingly, more infant siblings who developed autism oriented towards their names at 4 months of age than infant siblings who did not develop autism (Yirmiya et al., 2006). Similarly, Onzonoff et al. (2010) found that infants who developed autism exhibited non-significantly more social attention at 6 months but by 12 months displayed less social attention than developmentally delayed and typically developing infants. Decreases in social engagement between 6 and 12 months were also documented in case-studies of a sub-group of infants who developed autism (Bryson et al., 2007). These findings of decreases in social attention between 6 and 12 months in prospective designs are in contrast with those observed by Maestro et al. (2005) with a retrospective design.

A slower rate of RJA development between 15 and 34 months of age is also associated with autism diagnoses (Yoder et al., 2009), as is reduced RJA between 12 and 24 months (Rozga et al., 2011; Landa et al., 2007, Sullivan et al., 2007; Yoder et al., 2009). Decreased eye-contact (Goldberg et al., 2005) and social referencing (Cornew, Dobkins, Akshoomoff, McCleery, & Carver, in press) at around 18 months are also associated with autism.

Social Responsiveness and the Broader Autism Phenotype in Infancy

Reduced eye contact (Goldberg et al., 2005) and decreased orienting to name may also be a characteristic of the broader autism phenotype (Nadig et al., 2007). Orienting to name was associated with subsequent cases of developmental abnormalities, including autism, among infant siblings but not among low-risk infants.

Whether decreased RJA is a characteristic of the broader autism phenotype is less clear. While some comparisons of infant siblings to low-risk infants documented reduced RJA among infant siblings (Cassell et al., 2007; Presmanes, Walden, Stone, & Yoder, 2007) others did not
(Goldberg et al., 2005; Toth, Dawson, Meltzoff, Greenson, & Fein, 2007, Yirmiya et al., 2006). Neither of the studies yielding RJA differences between infant siblings and low-risk infants distinguished between infant siblings who did and did not develop autism. Thus, reduced RJA among infant siblings may have been driven by the subgroup of infant siblings who went on to develop autism. Studies which did compare infant siblings who did not become autistic to low-risk infants found no evidence of RJA deficits in non-autistic infant siblings (Rozga et al., 2011; Toth et al., 2007; Yirmiya et al., 2006).

While these findings suggests that RJA deficits are specific to autistic children rather than an aspect of the broader autism phenotype, only a subset of the relatives of autistic individuals display subclinical characteristics of autism. One study examined the development of RJA in infant siblings who became autistic, infant siblings with subclinical symptoms, and typically developing infant siblings (Sullivan et al., 2007). Only the infants who went on to develop autism exhibited significantly reduced RJA at 24 months of age.

Thus, atypical social attention, but perhaps not reduced RJA, by 12 months of age is probably a characteristic of the broader autism phenotype. A number of other studies have documented divergent responses to social stimuli between infant siblings and low-risk infants by 6 months of age (e.g. Bhat, Galloway, & Landa, 2010; Noland, Reznick, Stone, Walden, & Sheridan, 2010). Because diagnostic outcomes were often not ascertained in these studies, it is difficult to discern if observed differences are a characteristic of the broader autism phenotype or due to a subgroup of infants who later developed autism.

**Current Study**
As discussed above, few clear behavioral markers of autism and of the broader autism phenotype have been identified prospectively prior to 12 months of age. However, neural evidence of reduced responsiveness to dynamic gaze shifts is evident among 6 to 10 month olds who develop autism (Elsabbagh et al., 2012). Thus, subtle differences between infants who do and do not develop autism may be apparent with more precise measures.

Because eye-tracking enhances the detection of subtle shifts in visual attention (Aslin, 2007), it is a promising tool for identifying subtle behavioral differences between infants who do and do not develop autism. An eye-tracking assessment of toddlers with and without autism revealed that toddlers on the spectrum only differed from controls in their interest in a televised model when the model directed dyadic cues in their direction and not when she was looking elsewhere (Chawarska et al., in press). The current study is one of the first (none yet published) to use eye-tracking to determine if reduced interest in dyadic cues (direct gaze coupled with child-directed speech) is predictive of subsequent autism diagnoses.

Eye-tracking studies of gaze following have been conducted with typically developing infants (Gredebäck, Theuring, Hauf, & Kenward, 2008; Gredebäck, Fikke, & Melinder, 2010; Senju & Csibra, 2008; von Hofsten, Dahlström, & Fredriksson, 2005) and infant siblings (Bedford et al., in press). Eye-tracking and in-person measures of gaze following were found to be related among 18 month old infant siblings (Navab, Gillespie-Lynch, Sigman, Johnson, & Hutman, 2011). Unexpectedly given prospective (Rozga et al., 2011) and retrospective (Clifford & Dissanayake, 2008) reports that reduced RJA may be apparent by 12 months of age among infants who are later diagnosed with autism, infant siblings who were diagnosed with autism did not exhibit reduced gaze following in response to an eye-tracking assessment at 7 or 13 months of age (Bedford et al., in press). However, infant siblings with autism and other types of
developmental delays attended less to the object the model was attending to relative to everything else on the screen than typically developing infants. The authors interpreted this as evidence for decreased understanding of the referential significance of gaze (e.g. Brooks & Meltzoff, 2008).

Both the current study and the Bedford et al. study used stimuli adapted from a study documenting enhanced gaze following among developing infants following dyadic cues (Senju and Csibra, 2008). More dyadic cues were provided prior to each gaze following opportunity in the current study than in the other two studies in order to maximize opportunities for potential atypicalities of social attention to become apparent. Prior to the opportunity for gaze following, the model looked toward the infant in both this study and the Bedford et al. study but she did not smile excitedly nor speak in the Bedford et al. study. Interest in the model during this “social greeting” is one of the key variables in the current study and one of the exclusion criteria for a usable trial in the Bedford et al. study.

Hypotheses.

We expected reduced social attention to be both an early risk marker of autism and a characteristic of the broader autism phenotype. Thus, we expected that infants who developed autism would display less social attention than infants with subclinical characteristics of autism who in turn would show less social attention than typically developing infant siblings. Because even infant siblings who are classified as typically developing may exhibit some characteristics of the broader autism phenotype, we also investigated whether social differences that might potentially index the broader autism phenotype were apparent when comparing high-risk infants to low-risk infants. Comparisons based on risk were conducted irrespective of diagnostic
outcome because while we enrolled a low-risk comparison group, we did not obtain diagnostic outcomes for the low-risk group.

As the Bedford et al. (in press) findings had not been published when we were designing our study, we expected RJA in the eye-tracker to be predictive of subsequent autism diagnoses. We did not expect RJA to be a characteristic of the broader autism phenotype. Thus, we expected infant siblings who developed autism to demonstrate less RJA than those who did not develop autism. However, we did not expect infant siblings with subclinical characteristics to differ from typically developing infant sibs nor did we expect infant siblings to differ from low-risk infants.

**Methods**

**Participants**

**Infant siblings.**

Infant sibling participants were included based on their siblings' diagnosis with autistic disorder, confirmed by the UCLA Autism Evaluation Clinic. Confirmation of proband diagnosis was based on DSM-IV criteria (APA, 2000), the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000), and the Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, & LeCouteur 1994). Infant siblings were recruited through the UCLA Autism Evaluation Clinic, through other studies at the UCLA Center for Autism Research and Treatment, and through organizations that provide services for children with autism and their families.
The current study began partway through an ongoing longitudinal study wherein infants were assessed at 6, 12, 18, 24 and 36 months of age. Assessments were generally conducted within 2 weeks of a target age. Infant siblings were enrolled in the current study at 6 (N= 39) or 12 (N= 31), 18 (N= 27), or 24 (N= 7) months of age. They viewed the same eye-tracking stimulus at subsequent visits between 12 and 24 months of age. See Table 1 for the number of participants at each time point who provided usable data for whom it was a first, second, third or fourth visit.

**Diagnostic outcome.**

At the 36 month visit, infant siblings were classified into one of three groups: an autism spectrum condition group (ASC), a group of participants who exhibited non-specific developmental concerns (Concerns), and those who did not exhibit concerns (TD: Typically Developing). An ASC diagnosis (according to DSM-IV criteria) was conferred by a clinician on the basis of ADOS scores. The Social Communication Questionnaire was used to verify that symptoms were also present in the home (SCQ: Rutter, Bailey, Lord, & Berument, 2003). The Concerns group was defined by elevated ADOS scores (within one point of the cutoff for ASC) or developmental delays indicated by scores two standard deviations below the mean on one of the scales of the Mullen Scales of Early Learning (MSEL, Mullen, 1995) or scores one standard

<table>
<thead>
<tr>
<th></th>
<th>Low-Risk</th>
<th></th>
<th>High-Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
</tr>
<tr>
<td>6</td>
<td>23</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>13</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>10</td>
<td>13</td>
<td>4</td>
</tr>
</tbody>
</table>

*Table 1. Number of participants for whom it was their 1st, 2nd, 3rd or 4th visit at each age (mos.)*
deviation below the mean on two or more MSEL scales. Participants who did not meet these criteria but were judged as concerning according to clinician’s best estimates (e.g., difficulty with pragmatics) were also included in the Concerns group. The TD group included all participants who neither met criteria for ASC nor Concerns. Participants (N = 2) who met criteria for ASC on the ADOS and received a clinical diagnosis at 24 months were classified as ASC for analyses. Participants who neither met criteria at 2 years nor yet were evaluated at 3 years were not assigned a diagnostic outcome for the following analyses (N=26).

Why did we wait until 3 years of age to assign diagnostic outcomes (except in the rare cases when children met ADOS criteria for autism and were diagnosed as autistic by a clinician at 2 years of age)? Substantial variability in the development of children with autism is often apparent prior to 3 years of age (Lord, Luyster, Guthrie, & Pickles, in press). Indeed in our sample, diagnostic outcome was not associated with ADOS severity scores (Gotham, Pickles, & Lord, 2009) at 24 months (p = .072) after controlling for gender (see Table 2). However, diagnostic outcome was associated with ADOS severity scores at 3 years (p = .004).

<table>
<thead>
<tr>
<th></th>
<th>Infant Sib TD</th>
<th>Concerns</th>
<th>ASC (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>1.69 (.105)</td>
<td>1.71 (1.25)</td>
<td>2.33 (1.73)</td>
</tr>
<tr>
<td>24</td>
<td>1.21 (.415)</td>
<td>2.57 (1.62)</td>
<td>3.45 (1.86)</td>
</tr>
<tr>
<td>36</td>
<td>1.50 (.688)</td>
<td>3.00 (2.00)</td>
<td>6.00 (1.63)</td>
</tr>
</tbody>
</table>

*Table 2. ADOS Severity Scores by diagnostic outcome Mean (SE)*

See Table 3 for the number of participants of each gender who participated at each time point in relation to risk group or diagnostic outcome. The number of high-risk infants at each visit exceeds the number of high-risk infants classified as ASC, other concerns or typically developing because diagnostic outcome is not yet available for a number of participants who provided data at 24 months of age.
Table 3. Gender composition of data (female: male) by risk or outcome.

<table>
<thead>
<tr>
<th></th>
<th>Low-Risk</th>
<th>High-Risk</th>
<th>HR-TD</th>
<th>HR-Concerns</th>
<th>HR-ASC</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>12:11</td>
<td>6:10</td>
<td>4:2</td>
<td>0:2</td>
<td>0:2</td>
</tr>
<tr>
<td>12</td>
<td>10:13</td>
<td>14:15</td>
<td>7:5</td>
<td>2:1</td>
<td>1:3</td>
</tr>
<tr>
<td>18</td>
<td>12:9</td>
<td>17:24</td>
<td>11:7</td>
<td>2:2</td>
<td>1:5</td>
</tr>
<tr>
<td>24</td>
<td>20:19</td>
<td>12:21</td>
<td>5:5</td>
<td>1:4</td>
<td>2:6</td>
</tr>
</tbody>
</table>

Low-risk infants.

Low-risk infants were recruited by letter and telephone through a commercial list of new parents. Low-risk participants were required to not have a family history of autism but they were not required to have an older sibling. Low-risk infants enrolled in the study at 6 (N= 35), 12 (N= 29), 18 (N= 24), or 24 (N= 12) months of age. While some low-risk infants participated in the study more than once, many participated only once (see Table 1). Neither diagnostic outcomes nor measures of developmental level were ascertained for low-risk infants. Assessments of infant siblings and low-risk infants occurred in two different laboratories although an identical type of eye tracker was used in both labs. Due to our lack of knowledge about diagnostic outcomes and the confounding of lab location with risk, comparisons between infant siblings and low-risk infants should be interpreted with caution.

Exclusions.

Infant siblings who did not return for either the 24 month or the 36 month visit (for whom neither preliminary symptomatology nor final diagnostic outcomes could be ascertained) were excluded from all analyses (N= 23). Fifteen low-risk infants were excluded after parents reported a family history of autism. Eighteen and twenty four month olds (N= 4) who scored as at elevated autism risk on the Modified Checklist for Autism in Toddlers (M-CHAT) were also excluded. Additionally, data from a given visit was excluded if eye-tracking quality was poor
due to excessive motion or poor calibrations or if the eye-tracking assessment was not completed
due to infant fussiness. Reliability of exclusions based on data quality was confirmed by
comparing a second coder’s evaluations of data quality to the primary researcher’s data quality
decisions (Cohen’s Kappa=.72 for 30% of the sample). Data lost due to computer malfunction
and experimenter error (incorrect screen resolution) were also excluded from analyses. See Table
4 for the number of participants who were excluded and why at each time point.

<table>
<thead>
<tr>
<th></th>
<th>Low-Risk</th>
<th></th>
<th>High-Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DQ</td>
<td>FA</td>
<td>DL</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>7</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

*Table 4.* Number of participants excluded at each age point by reason for exclusion. Participants
who were excluded from the study overall are not included in this table. DQ: Data quality
(excessive motion or poor calibration), FA: Failed administration due to infant fussiness, DL:
Data lost due to computer malfunction, WR: Data not usable due to wrong screen resolution

**Standardized Measures**

*Administered to infant siblings.*

*Mullen Scales of Early Learning (MSEL; Mullen, 1995).* The MSEL is a standardized
developmental assessment of cognitive and motor development. It measures verbal and non-
verbal IQ for children less than 6 years of age. It provides an overall index score as well as
verbal subscale scores (Receptive Language and Expressive Language) and non-verbal subscale
scores (Visual Reception and Fine Motor). The Mullen provides T scores, age equivalent scores,
and raw scores. The Mullen has good test-retest reliability and high internal consistency.
Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000). The ADOS is a semistructured, standardized observational assessment of social interaction, communication, play, and imaginative use of materials used to diagnose autism spectrum disorders. Different modules of the ADOS are administered based on a child’s language ability. Social-affective symptoms, restricted and repetitive behaviors and severity scores can be derived based upon the revised ADOS algorithms (Gotham et al., 2009). Higher ADOS severity scores indicate greater levels of autistic symptomatology.

Social Communication Questionnaire (SCQ: Rutter et al., 2003). The SCQ is a 40 item parent questionnaire which assesses a child’s developmental history and current functioning in order to identify individuals who may have autism.

Administered to low-risk infant and infant siblings.

Modified Checklist for Autism in Toddlers (M-CHAT; Robins et al., 2001). The M-Chat is a 23 item parent questionnaire used to assess autism risk in toddlers. Items assess social relatedness, communication, and atypical sensory behaviors. A child is classified as passing or failing the M-CHAT based on the number of potential symptoms of autism observed. Children who fail the M-CHAT are more likely to develop autism than those who do not.

Eye-tracking Assessment of Social attention and Gaze Following

Infant looking behaviors were recorded by a Tobii 1750 Eye Tracker, integrated with a 17-inch monitor, while the infant was seated on a parent’s lap approximately 65 cm from the monitor. Cameras beneath the monitor recorded reflections from an infrared light at a frequency of 50 Hz to assess the distance between the cornea and the pupil of both eyes. The accuracy of these recordings approximates .5-1° of visual angle. While the eye tracker compensates for head
movements, movements faster than 10 cm/s occasion 100-ms recovery time. Stimuli were displayed with ClearView software (Tobii Technology AB; www.tobii.com). Fixations were defined as gaze within a 30 pixel radius for at least 100 ms. The “normal” ClearView validity filter averaging across both eyes was used. A five-point calibration was administered prior to the assessment.

The eye-tracking gaze social attention and following stimuli was a modification of stimuli developed by Senju and Csibra (2008). Each of the opportunities for social attention and gaze following were preceded by a colorful, sound-paired, animated “attention-getter” that was displayed until the infant looked to the center of the screen. Once attention was secured, the pre-recorded stimuli replaced the attention getter. The video consisted of a black background and a model wearing a neutral-colored shirt with her hair tied back. Two colorful, identical Lego structures were placed in front and on either side of the model, atop a black table (see Figure 1).

![Attention Getter Baseline](image-url)
During the baseline period, the model’s gaze remained fixed on the table in front of her (~2 seconds). This phase was followed by a social greeting phase (~1.8 seconds), during which the model looked into the camera, smiled, and said in infant-directed speech: “Hello there.” The final stage, wherein the model turned her head toward one of the two objects and then fixated on the object, provided an opportunity for gaze following (~4 seconds). The model maintained a neutral facial expression and remained silent when turning her head and gazing at the object. Across the four trials presented to each child, the model attended twice to the object on her right and twice to the object on her left. The order of looks to either side was counterbalanced across participants. The stimuli developed by Senju and Csibra differed from the current study in two key ways: the model remained neutral throughout the stimuli and the objects were not identical in Senjun and Csibra’s stimuli.

The model’s face measured 5.1° and 3.6° of vertical and horizontal angle. Each object measured 2.3° and 2.9° of vertical and horizontal angle. Rectangular areas of interest (AOIs),
defined manually using Clearview software, subtended approximately 1° from the edge of stimuli.

**Eye-tracking variables.**

*Overall attention to stimuli.* (ms) This measure was designed to assess potential differences in overall attention to the stimuli. It was calculated by summing the duration the infant attended anywhere on the screen across all trials.

*Social attention score.* (ms) Building upon Chawarska et al.’s (in press) finding that toddlers with autism attend less than typically developing children to a face emitting dyadic cues but not to a face without dyadic cues, we calculated an *interest in dyadic cues* difference score by subtracting attention to the model’s head during baseline from attention to the model’s head during the social greeting. We wished to assess social attention independently of where a child was looking before or after the opportunity for social attention so as not to confound social attention with the tendency to orient to an attention getter or difficulty disengaging from a central stimulus. Difficulties with disengagement may be a characteristic of autism in infancy (e.g. Zwaigenbaum et al., 2005). Thus, all trials were considered valid for social attention variables as long as the child did not fuss out before all four trials were administered.

*Joint attention variables.*

Only fixations upon an object immediately preceded by a fixation upon the model’s face were considered for classification as successful or unsuccessful instances of gaze following. Higher ratios indicate higher levels of gaze following. Because our operational definition of gaze following required moving from the center of the screen (the head) to one of the objects during
the target phase, a usable trial was defined by at least one fixation upon the attention getter prior to each trial and one fixation upon the screen during the opportunity for gaze following.

A standard difference score was calculated by subtracting the frequency with which the infant’s first look from the model to an object was incongruent with the model’s gaze from the frequency with which the first look was congruent (Gredebäck et al., 2010 based upon Corkum & Moore, 1998).

A duration difference score was calculated by subtracting the overall amount of time the infant attended to the objects during incorrect first looks from the total amount of time attended during correct first looks.

Results

Analytic Approach

The results reported in this dissertation remain preliminary as diagnostic outcomes are not yet available for 26 infant siblings who had not dropped from the study as of 24 months of age. The joint attention variables were log transformed because they exhibited excessive skew and/or kurtosis. ANOVAs were used to assess relations between eye-tracking data at 6 or 12 months of age (depending when infants provided their first usable data points) and diagnostic outcome or risk group. Linear mixed models were then used to analyze relations between eye-tracking data across the course of the study (all data from a given child could be used in such analyses) and outcomes.

Effectiveness of Eye-Tracking Measure of Gaze Following

Before focusing in on our main research questions, it was important to determine if our eye-tracking measure was effective at eliciting gaze following across development. Gaze following typically does not become fully developed until 18 months of age (Butler, Caron &
Brooks, 2009; Butterworth & Jarrett, 1991; Deák, Flom, & Pick, 2000). Bedford et al. (in press) found that infants’ first looks were more frequently to the correct relative to the incorrect object at both 7 and 13 months of age. However, their stimuli differed from ours in four key ways: they used more trials (12 rather than 4), they used non-identical objects, they required infants to attend to the social greeting in order for data to be analyzed and the model’s expression was consistently neutral across the stimuli.

Did our stimuli elicit gaze following as effectively as theirs did? One sample t-tests ranks tests suggest that it did not. While both low-risk and high-risk infants looked more frequently to the correct relative to the incorrect object at both 18 and 24 months of age ($p < .001$), neither risk group looked more frequently to the correct relative to the incorrect object at 6 months of age ($p > .05$). Low-risk infants ($p = .042$) but not high-risk infants exhibited above chance gaze following by 12 months of age. Either differences in selection criteria for usable trials, less trials, identical objects (associated with less gaze following in the interactive literature e.g. Deak et al., 2000) or a shift in apparent mood between the social greeting (happy) and the opportunity for gaze following (neutral) contributed to less consistent gaze following at 6 months of age in the current study relative to the Bedford et al study.

**Relations between Eye-tracking in Infancy and Diagnostic Outcomes**

The following ANOVAs contain age, gender, diagnostic classification and the interaction between diagnostic classification and age. Given the limited power of the current analyses, it is important to note that the same patterns emerge if only diagnostic classification and age are included in analyses. The interaction between diagnostic classification and gender was not included because there is only one female in the autism group at 6 and 12 months of age (see Table 3).
No relations between diagnostic outcome and overall attention to stimuli were observed ($p=.636$). However, infants who developed autism ($M = -289.14, SE = 812.86$) exhibited less interest in dyadic cues than typically developing infants ($M = 3196.53, SE = 484.81$; overall model $F (1, 18) = 6.295; p = .008, R^2 = .318$; post-hoc $p = .007$; see Figure 2). Infants who developed other concerns ($M = 1857.33, SE = 806.18$) did not differ from either of the other two groups.

![Figure 2. Relations between Interest in Dyadic Cues in Infancy and Diagnostic Outcomes](image)

Neither the standard difference score ($p = .971$; see Figure 3) nor the duration difference score ($p = .789$) varied by diagnosis.
We had chosen to use a difference score as our measure of social attention (i.e. attention to the model emitting dyadic cues versus when she was looking down) to build upon Chawarska et al’s (in press) finding but also because attention during the social greeting exhibited excessive skew and kurtosis that did not improve with data transformations. Skew and kurtosis were not apparent after subtracting attention to the model during baseline from attention to the model during the social greeting.

In order to examine whether differences in interest in dyadic cues scores between infants with and without autism were due to greater looking time to dyadic cues relative to baseline for typically developing infants but not those with autism, we conducted repeated measures analyses with z-scores of attention to the model during baseline and during the social greeting as dependent measures. These analyses should be interpreted with caution because attention to the social greeting was not normally distributed. They revealed that typically developing infants ($p=.005$), but not those with autism ($p=.287$) or other concerns ($p=.469$), attended more to the
model during the social greeting than during baseline. Thus, the lack of enhanced attention to a face emitting dyadic cues relative to a face looking down distinguished infants who developed autism from those who did not while gaze following did not.

**Relations between Eye-tracking in Infancy and Risk Group**

In order to investigate possible aspects of the broader autism phenotype in infancy, the same set of predictors described above were entered into ANOVAs except that diagnostic outcome was replaced with risk group. Low-risk and high-risk infants did not differ in terms of any of the eye-tracking variables except for overall attention to stimuli (\( F(7, 74) = 2.577; p = .020 \), adjusted \( R^2 = .120 \)). Overall duration of attending to the stimuli was greater among low-risk infants (\( M = 33166.65, SE = 947.349 \)) than high-risk infants (\( M = 28728.662, SE = 1069.131 \)). This difference should be interpreted with caution as it could be due to the different laboratories in which the two risk groups participated (i.e. the low-risk group was run in a dark lab with black curtains while the infants siblings participated in a lighter room where the eye-tracker was surrounded by a light blue wooden structure).

**Relations between Eye-tracking from 6 to 24 Months and Diagnostic Outcomes**

The following analyses use mixed models to examine all data from 6 to 24 months of age. Gender, diagnostic outcome and interactions between gender and diagnostic outcome and between age and diagnostic outcome were entered into a mixed linear model predicting each of the eye-tracking variables.

No relationship between outcome and overall attention to the stimuli was observed (\( p = .305 \)). Diagnostic outcome \( F(2, 65) = 4.266; p = .018 \), age \( F(1, 39) = 4.539; p = .039 \), and the interaction between age and outcome \( F(2, 40) = 4.878; p = .013 \) were associated with interest in dyadic cues. Individuals with ASC \( t(66.023) = -3.758, p = .013 \), but not those with other
concerns \( (p = .383) \), exhibited less social attention than typically developing participants. Social attention decreased with age overall \( t (49.376) = -3.758, p = .013 \) but increased with age for individuals on the spectrum \( t (42.304) = 3.033, p = .004 \) relative to typically developing individuals. No relations between any of the predictors and the standard difference score were observed \( (p = .951) \). Only age was positively associated with the duration difference score \( t (60.963) = 2.363, p = .021 \). Again, no reductions in gaze following were observed among infants with autism. However, infants with autism and older participants exhibited less interest in dyadic cues than their counterparts. Interest in dyadic cues increased with age among individuals with autism relative to typically developing toddlers.

**Discussion**

The current study demonstrates that reduced interest in dyadic cues in infancy and between 6 and 24 months of age is associated with subsequent autism diagnoses. It extends Chawarska et al.’s (in press) finding demonstrating reduced social attention among toddlers on the spectrum at a mean age of 23 months by suggesting that relative disinterest in dyadic cues may be an early risk marker of autism. Despite generally decreased attention to dyadic cues from 6 to 24 months of age, toddlers on the spectrum showed increased interest in dyadic cues with age relative to typically developing children. This progression from less to more social interest is more consistent with Maestro et al.’s (2005) retrospective research, albeit over a longer time frame, than with prospective research documenting increases in social difficulties between 6 and 12 months of age in autism (Bryson et al., 2007; Onzonoff et al. (2010).

Despite difficulties collecting longitudinal data even at the relatively sparse time frames used in this study, these results suggest that future research might benefit from multiple assessments between 0 and 24 months of age to really capture when decreased interest in dyadic
cues becomes apparent, how attention to dyadic cues changes with time, and potential relations between dyadic and triadic atypicalities in autism. Gaze contingent stimuli might be particularly useful for this type of frequently recurring assessment because it would be more likely to remain interesting across multiple presentations and is more similar to in-person social interactions than prerecorded stimuli. Given substantial variability in the early development of individuals with autism (e.g. Lord et al., in press), a better understanding of developmental trajectories of social attention associated with autism in infancy may assist in the identification of different endophenotypes associated with autism.

The current finding of decreased interest in dyadic cues among infants later diagnosed with autism is consistent with many retrospective and prospective studies documenting reduced orienting to name and attention to faces among infants and toddlers who go on to develop autism (e.g. Baranek, 1999; Hutman et al., 2010; Osterling & Dawson, 1994; Osterlig et al., 2002; Ozonoff et al., 2010; Maestro et al., 2005, Nadig et al., 2007; Werner et al., 2002; Zwaigenbaum et al., 2005). Because so few of our 6 month old participants have developed autism, it is not possible to state based on the current study if eye-tracking can identify infants who will go on to develop autism earlier than other prospective measures can. At the very least, eye-tracking identifies social difficulties associated with a subsequent autism diagnosis as early as prospective behavioral measures do.

The current results suggest that decreased interest in dyadic cues may also be a characteristic of the broader autism phenotype. Infants with other developmental concerns exhibited intermediate levels of interest in dyadic cues that did not differ from either typically developing infants or those with autism. Like infants with autism, and unlike typically developing infants, they did not show enhanced attention to dyadic cues. However, the absence
of enhanced attention to dyadic cues may also be attributable to power issues as fewer infants
developed autism or other concerns relative to the number of infants who were typically
developing. Future research with larger sample sizes should compare eye-tracking measures of
social attention to in-person measures of social attention. Future research should evaluate
different types of dyadic cues separately (i.e. smiling, child-directed speech and direct gaze) as
well as in combination to discern if reduced responsiveness to certain types of ostensive cues is
related to autism in particular while reduced responsiveness to other types of ostensive cues is
associated with the broader autism phenotype more generally. Control conditions wherein similar
types of movements and sounds are apparent (e.g. Chawarska et al., in press) but dyadic cues are
not present would allow researchers to determine if reduced interest in dyadic cues is due to
social or non-social aspects of the stimuli.

No evidence that reduced gaze following in infancy is a marker of autism was apparent in
both the current study and the study by Bedford et al. (in press). Only one prospective study
(Rozga et al., 2011) and one retrospective study (Clifford & Dissanayake, 2008) documented
reduced RJA by 12 months among individuals who were later diagnosed with autism. Whether
RJA is commonly impaired by 12 months among infants who develop autism remains an open
question. Presmanes, Walden, Stone and Yoder (2007) found that moderately redundant RJA
cues discriminated better between infant siblings and low-risk infants than very redundant or
very sparse cues. Future studies of gaze following among infant siblings should vary the
redundancy of cues, including very sparse measures such as assessments of reflexive gaze
following.

The current study and the study by Bedford et al. (in press) suggest that the particular
type of eye-tracking measure of gaze following used in both studies (i.e. non-contingent with
multiple overlapping cues) is probably not the best early marker of autism. While gaze following as assessed with the current eye-tracking stimuli was related to in-person measures of RJA among 18 month old infant siblings (Navab et al., 2011), relatively weak relations between gaze following and age were observed in the current study. Responsiveness to interactive measures of RJA often increases dramatically from 6 to 18 months of age (Butler, Caron & Brooks, 2009; Butterworth & Jarrett, 1991; Deák, Flom, & Pick, 2000). The lack of strong changes in RJA in response to the current stimuli with age suggest that there may be a difference between RJA as assessed in the eye-tracker and in person. One possible difference is the lack of contingency in pre-recorded measures of gaze following. Another possibility is a limitation of the current stimuli, the shift between a big smile during the social greeting and the “neutral” face during the opportunity for gaze following often led parents to ask why the model looked so grumpy while looking at the objects. Older children often waved in response to the model’s greeting over the first few trials and then shrank back against their parents when not receiving a response. Thus, future studies should keep facial expression consistent and perhaps do away with a greeting that leads to no further “social interaction”.

However, neither the apparent shift in mood nor the greeting occurred in the Bedford et al. (in press) stimuli and they too did not observe reduced gaze following among infants who went on to develop autism. In two eye-tracing studies examining RJA among typically infants, Gredebäck et al. (2008; 2010) found evidence that greater RJA occurs in response to interactive eye-tracking measures than pre-recorded stimuli. In order to benefit from the sensitivity of eye-tracking while eliciting greater levels of RJA than observed in response to the pre-recorded stimuli in the current study, future studies should use interactive eye-tracking measures of RJA and social attention to compare infants at high- and low- risk for autism.
The difference in overall attention to the stimuli between infant siblings and low-risk infants, in conjunction with few other differences between the two groups, suggests that future comparisons of low-risk and high-risk infants should be conducted in the same laboratory in order to rule out context based reasons for differences in attentiveness.

Despite the limitations of the current study, it suggests promising avenues for future research aimed at identifying very early markers of autism and the broader autism phenotype. While atypicalities of social responsiveness are one of the defining characteristics of autism across development, operationalizing “social” can be a surprisingly difficult enterprise. The current study suggests that it may not just be the presence or absence of people in images that makes a scene social but rather the presence or absence of dyadic cues. Joint attention is believed to grow from affectively charged dyadic interactions (Adamson & Bakeman, 1985; Kasari, Sigman, Mundy & Yirmiya, 1990). Given relations between attention to dyadic cues in infancy and subsequent diagnoses in the absence of atypicalities of gaze following in response to redundant cues in the current study, it might be particularly fruitful to assess both attention to dyadic cues and low level gaze cueing among infant siblings and low-risk infants in order to determine if dyadic atypicalities precede or are preceded by impairments in reflexive gaze processing in autism. Investigations of the developmental relationship between dyadic attention and reflexive gaze following in infancy could elucidate a key question in autism research: is the development of joint attention atypical or simply delayed in autism?
Study 3: Atypical Gaze Following in Autism: A Comparison of Three Potential Mechanisms

Abstract

Atypical gaze following has long-term implications for the development of individuals with autism. Previous research has yielded conflicting evidence for mechanisms underlying atypical gaze following, and these contradictory findings may be partially attributable to the use of varied paradigms to assess gaze following at different developmental levels. To address this possibility, we administered the same set of paradigms to young children with autism ($N=21$) and chronological ($N=21$) and nonverbal mental age ($N=21$) matched controls, evaluating three potential mechanisms underlying atypical gaze following in autism: impaired reflexive gaze following, difficulty integrating gaze and affect, and reduced understanding of the referential significance of gaze. Children with autism exhibited impaired reflexive gaze following. No evidence of integration of gaze and affect, regardless of diagnosis, was observed. Reduced higher-level gaze following was apparent among children with autism during both an eye-tracking and in-person assessment. Word learning from social cues was better explained by developmental level than autism. Taken together, these results suggest that both high- and low-level atypicalities of gaze following may be a core characteristic of autism early in development and that referential use of gaze may emerge with development despite reduced gaze following. Thus, gaze following may traverse an atypical, rather than just delayed, trajectory in autism.
Under what conditions is responsiveness to joint attention (RJA), defined behaviorally as gaze or point following, reduced among young children with autism? Conflicting accounts have arisen from studies examining different aspects of RJA in autism (see Nation & Penny, 2008 for a review). Given the heterogeneity of autism itself (e.g. Geschwind, 2009), a reasonable way to resolve conflicting findings may be to administer a range of related tests to the same individuals. Such an approach yielded intriguing evidence that the ability to report where someone is looking may be intact among children with autism who do not demonstrate spontaneous RJA (Leekam, Baron-Cohen, Perrett, Milders & Brown, 1997). However, the current study is the first, to our knowledge, to examine three potential underlying mechanisms of atypical RJA in autism among the same individuals: (a) atypical orienting to low-level gaze cues (Goldberg et al., 2008; Johnson et al., 2005; Ristic et al., 2005), (b) difficulty integrating gaze and affect (de Jong, van Engeland & Kemner, 2008; Kasari, Sigman, Mundy & Yirmiya, 1990; Uono, Sato & Toichi, 2009), and (c) reduced recognition of the referential significance of gaze (Baron-Cohen, Baldwin, & Crowson, 1997; Preissler & Carey, 2005). These potential mechanisms are not necessarily independent and all could arise at least partially through reinforcement learning, which has been shown to play a role in RJA development (Corkum & Moore, 1998).

RJA typically begins to emerge very early in infancy (D’Entremont, Hains & Muir, 1997; Farroni, Massaccesi, Pividori, & Johnson, 2004; Hood, Willen, & Driver, 1998; Scaife & Bruner, 1975) and develops gradually from birth to around 18 months of age (Butler, Caron & Brooks, 2009; Butterworth & Jarrett, 1991; Deák, Flom, & Pick, 2000). Often emerging in the context of affective interactions (Adamson & Bakeman, 1985; Teufel et al., 2010), RJA is an adaptive skill in that it provides information about what others find relevant in the environment. However, gaze following may not initially be synonymous with an understanding of others’ viewpoints (Corkum
RJA is not always observed in children with autism (e.g. Dawson et al., 2004; Loveland & Landry, 1986; Sigman & Ruskin, 1999). Indeed, reduced RJA is an early predictor of autism (Landa, Holman, & Garrett-Mayer, 2007; Rozga et al., 2011; Sullivan et al., 2007; Yoder, Stone, Walden, & Malesa, 2009). Among individuals with ASCs (autism spectrum conditions), individual differences in RJA are related to concurrent linguistic and cognitive skills (e.g. Leekam, Hunnisett, & Moore, 1998; Sigman & Ruskin, 1999) and subsequent cognitive (Sigman & Ruskin, 1999), social-communicative (Mundy, Sigman, & Kasari, 1990; Sigman & McGovern, 2005; Sigman & Ruskin, 1999) and adaptive development (Gillespie-Lynch et al., 2012).

Reduced RJA is not consistent across studies, however. For example, RJA deficits were not observed in children with ASC who had a nonverbal mental age (NVMA) above 19 months (Mundy, Sigman, & Kasari, 1994), a verbal mental age (VMA) above 47 months (Leekam et al., 1998; but see Leekam et al., 1997) or a nonverbal IQ in the normal range (Leekam, López, & Moore, 2000). A delay in the development of RJA is apparent even when comparing children with autism to children with other types of developmental delay (e.g. Dawson et al., 1998; Leekam et al., 2000; Sigman & Ruskin, 1999). Leekam et al. (2000) surmised that children with autism may require both more time and greater intellectual development in order to interpret the predictive meaning of gaze cues.

Given the discrepant findings regarding the nature of the developmental period during which RJA deficits may be apparent, it is important to consider developmental level when comparing across studies examining gaze following in autism. For example, most studies of
reflexive gaze following were conducted with high-functioning participants in middle childhood or older.

**Is Reflexive Gaze Following Impaired in Autism?**

Evidence of reflexive gaze following in response to schematic faces may be apparent within a few days of birth (Farroni, Massaccesi, Pividori, & Johnson, 2004). Reflexive gaze following is typically assessed using attention cueing paradigms wherein participants attend to a central face presenting gaze cues that are predictive, counter-predictive or sometimes unrelated to the future location of a target (Posner, 1980). A validity effect, or speeded detection of targets that are validly cued (in the location that the model’s eyes were looking toward) relative to those that are invalidly cued (in the location the eyes were looking away from), is taken as evidence of reflexive orienting (e.g. Langton, Watt, & Bruce, 2000). Participants are often instructed to attend to targets covertly (i.e. without moving their eyes from the center of the screen) and detection is generally indicated by pressing a button. The majority of the evidence suggests that high functioning children and adolescents (mean age of 9 to 11 years of age) and adults with autism demonstrate intact reflexive orienting to gaze when assessed with Posner-style paradigms (de Jong, van Engeland & Kemner, 2008; Kylliainen & Hietanen, 2004; Pruett et al., 2011 Rutherford & Kryska, 2008; Senju et al., 2004; Swettenham et al., 2003; Uono, Sato & Toichi, 2009; Vaidya, Foss-Feig, Shook, Kaplan, Kenworthy & Gaillard, 2011; Vlamings, Stauder, van Son & Mottron, 2003; but see Goldberg et al., 2008; Ristic et al., 2005). Even when reflexive gaze following is intact, however, people with autism may not privilege gaze cues relative to non-social cues as much as typically developing individuals do (Chawarska, Klin & Volkmar, 2003; Greene et al., 2011; Senju, Tojo, Dairoku & Hasegawa, 2004; Vlamings et al., 2003; but see Kuhn et al., 2010; Ristic et al., 2002).
The aforementioned studies assessed covert reflexive orienting in high functioning individuals who would not be expected to exhibit reduced spontaneous RJA. Covert orienting may be unimpaired relative to overt orienting in autism (Gernsbacher et al., 2008). Notably, little is known about overt reflexive gaze following in autism. Kuhn et al. (2010) reported that overt reflexive orienting was intact among high functioning adults with autism, but they did not assess spontaneous gaze following. Instead, participants were instructed to move their eyes as quickly as possible to the left or right based on a color cue at the center of the gaze cue.

Other assessments of overt spontaneous reflexive gaze cueing recorded the looking behaviors of young children with autism to examine whether preschoolers with autism do (Chawarska et al., 2003) or do not (Johnson et al., 2005) spontaneously orient to eye gaze cues. Chawarska et al. (2003) compared 15 cognitively delayed 2 year olds with autism to 12 controls matched for chronological age (CA). Although the children with autism exhibited reduced RJA during an in-person assessment, the interaction between reflexive gaze following and diagnosis was not statistically significant. However, children with autism attended to the target faster regardless of cue validity. Because no differences in reaction time were observed when the central cue was not social, the authors interpreted this as evidence for decreased engagement with a social stimulus. Johnson et al. (2005) found that 9 young children (mean age of 33 months) with autism also oriented more quickly to targets regardless of gaze cues relative to language delayed and typically developing controls, and post-hoc planned t-tests after a marginally significant interaction between group and validity revealed that only the children with autism did not exhibit reflexive gaze following. A primary aim of the current study, therefore, was to determine if overt reflexive gaze following is intact among young children with autism.

**Integrating Gaze and Affect**
The preponderance of the evidence suggests that reflexive gaze following may be unimpaired in autism, yet even high functioning people with autism may have difficulty integrating emotional signals and gaze cues (de Jong et al., 2008; Uono, Sato & Toichi, 2009). Attention cueing paradigms typically present gaze cues in the context of neutral facial expressions. When gaze cues are paired with emotional expressions, typically developing adults may demonstrate more of a covert validity effect to fearful relative to happy or neutral gaze cues (Friesen, Halvorson, & Graham, 2011; Graham, Friesen, Fichtenholtz, & LaBar, 2010; Pecchinenda, Pes, Ferlazzo, & Zoccolotti, 2008; Tipples, 2006). However, emotional faces may not enhance reflexive orienting among individuals with autism. For example, 11 high-functioning adolescents with ASCs differed from age and IQ matched controls in that they did not show an enhanced covert validity effect for fearful relative to neutral gaze cues (Uono et al., 2009). Similarly, high functioning school-age children with autism did not demonstrate differential ERP responses to emotional relative to neutral gaze cues (de Jong et al., 2008).

Previous studies examining difficulties integrating gaze and emotion in autism have focused on covert orienting by high-functioning individuals and on the distinction between fear and neutral emotions (de Jong et al., 2008; Uono et al., 2009). Because children with autism may integrate positive affect with joint attention less than typically developing children and children with other disabilities (Kasari et al., 1990), a second aim of the current study was to examine overt gaze cueing in response to happy, fearful, and neutral expressions.

**Referential Significance of Gaze**

Explanations of reduced RJA in autism discussed previously focused primarily on atypicalities of social perception and on the intersection between perception and affect. A third
possibility is that RJA impairments arise from social-cognitive difficulties such as difficulty interpreting the referential significance of gaze (Baron-Cohen, Leslie, & Frith, 1985). Intention reading is associated with RJA performance (Schietecatte, Roeyers & Warreyn, in press) among children with autism. Nevertheless, an understanding of mental states is not necessary for gaze following, though such an understanding could motivate attention to gaze cues (see Nation & Penny, 2008). Difficulties using RJA to learn words in the absence of perceptual and perceptual-affective deficits would yield support for the theory that RJA impairments in autism arise from difficulties understanding referential intention.

Although understanding reference is not necessary for gaze following, gaze following is important for recognizing the referential nature of vocalizations in mapping words to objects (Baldwin, 1991; Scaife & Bruner, 1975; Seibert, Sliwin, & Hogan, 1986). Difficulties using a speaker’s focus of attention to learn words have been documented among severely cognitively delayed children with autism (Baron-Cohen, Baldwin, & Crowson, 1997; Preissler & Carey, 2005). Unlike typically developing and intellectually delayed children, however, school age children with autism were more likely to use a listener’s direction of gaze (LDG) than a speaker’s direction of gaze (SDG) strategy for word learning (Baron-Cohen, Baldwin, & Crowson, 1997; Preissler & Carey, 2005). Both groups of children learned words when the examiner labeled the object they were looking at (the follow-in condition). When attending to a different object than the examiner (the discrepant condition), most children with autism differed from controls in that they attached a novel label to the object they were attending to (LDG) rather than the object the individual labeling the object was attending to (SDG).

Baron-Cohen et al. (1997) proposed that the LDG word learning strategy was part of a general difficulty sharing attention and understanding intentions in autism. Yet the LDG strategy
may not be as common among children with autism as early studies suggested (Akechi et al., 2011; Luyster & Lord, 2009; Parish-Morris et al., 2007). Luyster and Lord (2009) examined whether difficulty using social cues to learn words was apparent among younger (mean age 30 months) and higher functioning ((with a mean NVIQ of 95) children with autism who were not particularly impaired at gaze following as assessed during the Autism Diagnostic Observation Schedule (ADOS). Children with autism and expressive language matched controls were more likely to select the object they were attending to after follow-in labeling and less likely to select the object they were attending to after discrepant labeling—that is, neither group commonly used the LDG strategy. However, children with autism did not differ from chance in their likelihood of selecting the object the examiner was attending to during discrepant labeling. Even among the 5 severely linguistically delayed children in the sample, one used the SDG strategy and 4 did not select any object. A similar pattern of avoiding mapping errors (but not necessary correctly mapping) is observed among 12 to 18 month old typically developing infants (Baldwin, 1999; Hollich et al., 2004), leading the authors to suggest that children with autism may follow a delayed but qualitatively typical path toward learning words from gaze cues.

Similar evidence of a delayed ability to use others’ attention to learn words despite infrequent mapping errors was documented among slightly older (mean age of 5 years) moderately cognitively impaired children (with a mean IQ of 76) with autism in comparison to VMA and NVMA matched controls (Parish-Morris et al., 2007). Separate assessments were used to measure joint attention and word learning in response to social cues (the examiner pointed to or touched a perceptually interesting or a perceptually dull object while labeling it). Children with autism looked to the indicated object during joint attention opportunities more frequently than expected by chance but less frequently and for a shorter duration than controls. While they
learned the labels of “interesting” objects, they did not (unlike controls) learn the labels of “dull” objects.

These studies imply that the ability to learn words from gaze cues may be related to gaze following abilities, but gaze following during word learning was not assessed. A recent eye-tracking study suggests that subtle atypicalities in gaze following may underlie subtle difficulties with word learning among high functioning individuals (Akechi et al., 2011). Seventeen children with ASCs (mean age of 9 years) were compared to typically developing controls. Pairs of novel objects were presented on a computer screen while a schematic face looked toward and labeled either the object the child was attending to or the object the child was not attending to. Participants with and without ASCs more frequently looked from the face to the object the face was attending to relative to the other object. Typically developing participants, but not those with ASCs, differed from chance in the duration which they attended to the target relative to the other object (this analysis was a follow-up to a non-significant interaction between group and duration). No group differences in word comprehension were observed after follow-in training, but the ASC group performed slightly less well than the controls following discrepant labeling. However, both groups demonstrated above chance word comprehension even after discrepant training.

The Akechi et al. (2011) finding suggests that the duration of attending to the objects of gaze may be related to word learning. An eye-tracking study with 3-year-old siblings of children with autism suggests that understanding the referential intention behind gaze may be more important for word learning than gaze following (Gliga et al., in press). Despite the absence of clear differences in gaze following (either duration or frequency) toward the object a model was labeling, siblings with poorer social-communication skills looked less to the target object during
testing than typically developing children. Below chance preference for the attended object during training was always associated with below chance performance during test but the reverse was not the case.

This series of studies suggests that difficulties with reference are unlikely to underlie gaze following impairments in autism because gaze following seems to be one of many pathways to the understanding of reference. The third aim of the current study, therefore, was to clarify relations between gaze following and word learning in response to gaze cues using both an eye-tracker (video presentation) and a live model, all in relation to cognitive level.

**Hypotheses**

**Hypothesis 1: intact reflexive orienting.** We predicted that reflexive gaze cueing would be intact among children with autism and typically developing children. We expected a main effect of validity and no moderating effect of diagnosis.

**Hypothesis 2: impaired integration of gaze and affect.** We hypothesized that children with autism would be impaired in their ability to integrate gaze and affect. In particular, we predicted an interaction between diagnosis, emotion, and validity because typically developing children were expected to exhibit a larger validity effect to fearful relative to neutral or happy faces. In contrast, we predicted no emotional enhancement of the validity effect among children with autism.

**Hypothesis 3: difficulty with the referential significance of gaze is related to developmental level rather than autism.** We expected reduced gaze following to be related to diagnosis after controlling for differences in age or NVMA but word learning following gaze cues to be attributable to NVMA rather than diagnosis.
Methods

Participants. Twenty-four children with autism and 42 children without an autism diagnosis participated. Children with autism (range 2.4 to 6.7 years) were individually matched to typically developing children in terms of CA (within 3 months) or NVMA (within 6 months). Gender was matched except in one case of CA and one case of NVMA. CA and NVMA age control groups were not independent of one another in that higher-functioning participants with autism were often matched with the same individual for both CA and NVMA. Comparisons of CA matches and children with ASCs are presented in Table 1 while comparisons of NVMA and children with ASCs are presented in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>ASC (N=21)</th>
<th>CA Match (N=21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number Female</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>CA in months</td>
<td>56.48(14.81)</td>
<td>56.95(15.27)</td>
</tr>
<tr>
<td>NVMA in months</td>
<td>52.33(23.49)*</td>
<td>67.00(17.62)*</td>
</tr>
<tr>
<td>VMA in months</td>
<td>45.24(21.75)*</td>
<td>67.90(18.00)*</td>
</tr>
<tr>
<td>Therapy in hours</td>
<td>484.04(401.25)</td>
<td>NA</td>
</tr>
<tr>
<td>Maternal Ed. in years</td>
<td>16.40(1.77)</td>
<td>17.52(2.14)</td>
</tr>
<tr>
<td>% English in home</td>
<td>87.59(16.64)</td>
<td>89.50(22.12)</td>
</tr>
</tbody>
</table>

* = significant differences (p <.05)

Table 1. Participant characteristics of CA matches and participants with ASCs: Mean (SD)
Table 2. Participant characteristics of NVMA matches and participants with ASCs: Mean (SD)

<table>
<thead>
<tr>
<th></th>
<th>ASC (N=21)</th>
<th>NVMA Match(N=21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number Female</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>CA in months</td>
<td>56.48(14.81)*</td>
<td>44.29(19.65)*</td>
</tr>
<tr>
<td>NVMA in months</td>
<td>52.33(23.49)</td>
<td>52.33(23.26)</td>
</tr>
<tr>
<td>VMA in months</td>
<td>45.24(21.75)</td>
<td>51.72(22.06)</td>
</tr>
<tr>
<td>Therapy in hours</td>
<td>484.04(401.25)</td>
<td>NA</td>
</tr>
<tr>
<td>Maternal Ed. in years</td>
<td>16.40(1.77)</td>
<td>17.35(2.01)</td>
</tr>
<tr>
<td>% English in home</td>
<td>87.59(16.64)</td>
<td>86.33(22.45)</td>
</tr>
</tbody>
</table>

* = significant differences (\(p < .05\))

Most of the children with autism who participated in this study were recruited from families participating in an ongoing study of the longitudinal development of infant siblings of children with autism. In addition, 15 older siblings with autism, four at-risk siblings who developed autism, and four children with autism were recruited via fliers from the community. Two participants from the community had verified diagnoses of autism but no evidence that the children met criteria for ASC on the ADOS or the autism diagnostic interview (ADI-R); they were excluded from the analyses. For the remaining participants with ASCs (two female) clinical diagnoses were informed by the ADOS and sometimes the ADI-R, 12 met the ADOS cut-off for autism and nine met the ADOS cut-off for ASC.

Control participants were screened for autism symptoms with the Social Responsiveness Scale (SRS: if they were 4 years or older) and the Childhood Autism Rating Scale (CARS: if they were younger than 4). Two typically developing participants were excluded for elevated
symptoms on the SRS. Three controls were excluded for a family history of autism. One control was excluded because the eye-tracker failed to follow his gaze due to thick glasses. Four controls were excluded because they did not match a child with autism in NVMA or CA.

**Standardized Measures**

**CARS** (Schopler, Bourgondien, Wellman, & Love, 2010). The CARS includes a 36 item parent-report questionnaire that, in conjunction with examiner ratings, yields dimensional or categorical ratings of autism symptomatology. It has high internal consistency and inter-rater reliability.

**Differential Ability Scales II** (DAS: Elliot; 1990). The DAS can be used to assess nonverbal and verbal intelligence among children with a mental age between 2 years 6 months and 17 years 11 months (Elliot, 1990). It has well established internal and external reliability. The DAS was used with most participants who were mature enough to receive age equivalence scores on the measure. Younger or more delayed children were assessed with the Mullen Scales of Early Learning (MSEL).

**MSEL** (Mullen, 1995). The MSEL is a standardized measure of VMA and NVMA from birth to 68 months of age. It has good test-retest reliability and high internal consistency. Participants who appeared to be below 2 years 6 months in mental age were administered the MSEL. The four younger siblings of children with autism who themselves developed autism were assessed with the MSEL as was one of their matched controls. The convergent validity of the DAS and the MSEL has been established in autistic and non-autistic populations (Bishop, Guthrie, Coffing, & Lord, 2011).

**SRS** (Constantino, 2005). The SRS is a 65 item parent rating scale that assesses the
severity of autism symptoms. It yields a dimensional measure of autism symptomatology as well as categorical cut-offs believed to indicate atypically high levels of symptoms. It can be used with individuals 4 to 18 years of age and has high internal consistency and test-retest reliability.

**Gaze Following Assessments**

**Eye-tracking overview.** Looking behaviors were recorded by a Tobii 1750 eye-tracker (Tobii Technology AB, Falls Church VA) integrated with a 17-in. monitor, while the child sat approximately 65 cm. from the monitor. Cameras beneath the monitor recorded reflections from an infrared light at a frequency of 50 Hz to assess the distance between the cornea and the pupil of both eyes. The spatial accuracy of these recordings approximates .5-1° of visual angle. Stimuli were displayed with Tobii ClearView software. Fixations were defined as gaze within a 30 pixel radius for at least 100 ms The “normal” ClearView validity filter averaging across both eyes was used. A five-point calibration was administered prior to the assessment. Rectangular areas of interest (AOIs), defined with ClearView software, subtended approximately 1° from the edge of all stimuli.

Participants were assigned to view the attention cueing and joint attention word learning eye-tracking paradigms in counterbalanced orders. Controls were assigned to the same version and order of each type of stimuli as their match with autism whenever possible (this occurred more frequently with chronological matches because such matches were more predictable prior to testing). Prior to each trial in either eye-tracking paradigm, the child’s attention was centered whenever possible with an animated attention getter. Trials prior to which the child’s attention was not centered were excluded from analysis.

**Gaze cueing.** This paradigm was modeled after the study by Chawarska et al. (2003),
which also measured overt gaze orienting in young children with autism. We also included measures of facial expression and stimulus onset asynchrony (SOA), the amount of time between the gaze cue and target onset, of 167 and 400 ms (e.g., Friesen et al., 2011; Graham et al., 2010).

Figure 1. Gaze cueing stimuli

Each trial began with the model looking straight ahead with a neutral expression for 1000 ms (see Figure 1). This was followed by a static image of the model gazing to the left or right with one of the following facial expressions: happiness, fear, or a neutral expression. The model’s face measured 16.1° and 11.1° of vertical and horizontal angle. The model was selected from the Nim Stim catalogue (Tottenham et al., 2009) based on high validity ratings of the three facial expressions. A different model than the one used for the stimuli is portrayed in Figure 1 due to copyright issues. A usable trial was defined by attention to the eyes during the gaze cue. Then the face disappeared and was replaced by an object (either a green balloon or a red bottle of bubbles) to the left or right of where the face had been which remained on the screen for 1000 ms. The images of objects were selected from the Bank of Standardized Stimuli to equate them.
in terms of familiarity and object complexity (Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010). The objects measured 6.4° and 6.0° of vertical and horizontal angle.

The model’s gaze predicted the location of the target 50% of the time. There were 48 trials. Two randomly-ordered versions of the stimuli were generated to counterbalance location cued, validity, target, and target location with the constraint that no stimulus type could occur more than twice in a row. This order was then transformed by replacing all fearful expressions with happy expressions, all happy with neutral and all neutral with fearful in order to create the second version of the stimuli.

Assessments of joint attention word learning. Participants were provided two opportunities for word learning, one in-person and one with the eye tracker. The eye-tracking word learning opportunity always occurred before the in-person opportunity because young children may learn words better from screens when screen-based word learning is preceded by an interaction (Roseberry et al., 2009). Thus a strong order effect could have occurred if the order of conditions were counterbalanced. Four conditions were created in which object-word pairings were counterbalanced across in-person and eye-tracking conditions.

Two counterbalanced sets of training and test materials were assigned to either the in-person or eye-tracking opportunity for word learning. Within each set, two novel objects were always associated with the same two novel words during training. The novel words were monosyllabic yet phonologically distinct across languages to allow recruitment of children who were not monolingual. The two pairs of words were “bon” vs “dit” and “deet” vs “don.” During each training phase, children viewed a model seated between two novel objects. The objects always remained in the same location relative to the model. The model always turned in the
following order LRRLRLR and labeled each object once after she had turned to look at it. The word the model said as she looked to the left or right was counterbalanced (see Figure 2).

*Figure 2. Eye-tracking joint attention word learning stimuli*

**Eye-tracking assessment of joint attention word learning.** Prior to word learning training, the two novel objects were shown on the screen in silence for 2 s to assess a baseline preference for either object, then repeated with the location of the objects switched. This was followed by a training session consisting of 8 word learning opportunities (4 for each object). Each trial began with a baseline phase wherein the model looked down (~2 s) followed by gazing and smiling in the direction of the child (~1.8 s). Then she turned toward one of the novel objects, and labeled it once while gazing at it and continuing to smile (~5 s). The model’s face measured 7.9° and 5.3° of vertical and horizontal angle. The objects measured between 7.6° and 9.0° of vertical and 6.5° and 9.3° of horizontal angle.

The training phase was followed by a testing phase during which only the objects were visible on the screen (2 s). During each of these trials, the child heard the model say a trained
novel word once. Two same position trials (wherein the child heard each of the novel words while the location of the objects relative to one another was unchanged from training) were followed by two different position trials (wherein the location of the objects was reversed from training). This was followed by a “mutual exclusivity” trial wherein the child saw the two trained objects, a familiar object and a novel object that had not been seen before and heard a novel word that had not been heard before (i.e. “blick” or “bleek”). Lastly, two more same position and two more different position trials were presented. After the eye-tracking assessment, the child was presented with the two novel objects seen in the video and asked to select each novel word i.e. “Give me the deet.” Regardless of whether responses were correct, the examiner commended the child’s choice.

**In-person assessment of joint attention word learning.** This assessment was modeled after the eye-tracking measure. The child sat across the table from the model. The model secured the child’s attention before looking toward one of the objects while labeling it and smiling. After eight training trials, the child was asked to select each object in its trained position i.e. “Show me the deet”, followed by two reversal trials, a mutual exclusivity trial, and two more opportunities to select the objects in the same position as during training.

**Dependent variables**

*Gaze cueing usable trials.* Number of trials in which the child looked at the model’s eyes during the directional gaze cue and oriented to the target. If this was 0 for a given condition, no gaze cueing response time (RT) was available for that condition.

*Gaze cueing RT.* Average RT in ms to orient to the target following target appearance. Two RT were calculated for each pairing of emotion and SOA: i.e. valid RT (following a valid
gaze cue) and invalid RT.

*In-person first look difference score (Corkum & Moore, 1998).* Number of training trials in which the child first looked from the model’s head to the object the model was attending to minus the number of trials wherein the child first looked to the other object. Two independent coders (Cohen’s kappa .80) achieved reliability on 30% of the sample.

*Eye-tracking first look difference score.* This was the same as the variable previously described recorded with the eye-tracker.

*Eye-tracking referent preference (based on Hollich et al., 2000).* During labeling in test trials, the duration of looks at the correct referent relative to duration of attention to the incorrect object.

*Eye-tracking referent choice.* Children received a score of 1 if they selected the referent of both novel words and a score of 0 if they selected the referent of 1 or 0 words. All scores were assigned by the examiner during testing. This object selection was not video-taped so it was not possible for a second coder to double-check them.

*In-person referent choice.* Children were assigned a score of 1 if they selected the referent on 6 out of 6 test trials- i.e. better than chance- and 0 otherwise. Mutual exclusivity trials were not analyzed. All scores were assigned by the examiner during testing and double checked by a second coder.

Results

Despite careful attempts to individually match children with autism to controls, the number of children who provided usable data varied across measures. For each analysis, it will
be noted if the intended dimensions of matching still hold given loss of data. Because the CA and NVMA matched groups were not independent of one another, comparisons between each type of control and children with autism will be reported separately. CA matches are older on average than NVMA matches.

Response time variables were log transformed due to excessive skew. All significant main effects and interactions are reported. Repeated measures ANOVAs were conducted with diagnosis as a between subjects factor and measures of gaze following as repeated measures. When comparison groups differed in CA or NVMA, the variable that differed was entered into models as a covariate. Lower-level gaze following was operationalized as RT to valid and invalid cues. Higher-level gaze following was operationalized as first look differences scores during word learning opportunities in the eye-tracker and in-person. Binary logistic regressions were used to examine associations between NVMA, CA, diagnosis, first look difference scores and referent selection after word learning opportunities. An ANCOVA was used to examine relations between the same set of predictors and eye-tracking preference scores.

**Hypothesis 1: intact reflexive orienting.** Children with autism did not differ reliably from controls in the number of usable trials provided at either SOA. Among children who provided usable data at the short SOA, children with autism (N = 17) had marginally lower NVMA than CA matches (N = 18; p=.052) but did not differ reliably from NVMA matches (N = 17) in CA or NVMA. At the longer SOA, children with autism (N = 19) were significantly different than CA matches (N = 16) in NVMA (p = .017) but not NVMA matches (N = 15) in CA or NVMA.
Figure 3. Impaired reflexive gaze cueing at the 167 ms SOA for NVMA matches. Error bars represent SEs.

Focusing on CA matches, an ANCOVA with validity (valid RT and invalid RT) as a repeated measure, diagnosis as a between subjects factor, and NVMA as a covariate revealed no significant effects, although a trend toward an interaction between validity and diagnosis was observed ($p = .096$). An identical analysis conducted with NVMA matches (except without a covariate as the groups did not differ in CA or NVMA) revealed a main effect of diagnosis $F (1, 32) = 5.500$, $p = .025$ and an interaction between diagnosis and validity $F (1, 32) = 4.459$, $p = .043$ (see Figure 3). Post-hoc tests revealed a validity effect $F (1, 16) = 6.444$, $p = .022$ among typically developing participants but not children with autism. Children with ASCs ($M=2.494$, $SE= .020$) were also faster to orient to targets regardless of cue validity than controls ($M=2.559$, $SE= .020$; $p = .025$).

Focusing on CA matches at the longer SOA, an ANCOVA with validity, diagnosis and NVMA as a covariate revealed an effect of diagnosis, $F (1, 32) = 7.847$, $p = .003$ and an
interaction between validity and diagnosis $F(1, 32) = 4.391, p = .044$ (see Figure 4). Children with autism ($M=2.543, SE = .013$) were slower to orient to targets than typically developing children ($M=2.486, SD = .015; p = .009$). Post-hoc tests revealed a validity effect $F(1, 14) = 4.585, p = .050$ and a trend toward an interaction between validity and NVMA ($p = .060$) for typically developing children but not those with autism. No effects were observed when comparing NVMA matches.

![Figure 4](image-url)

*Figure 4.* Impaired gaze cueing at the 400 ms SOA for CA matches after controlling NVMA. Error bars represent $SE$s.

Contrary to our expectations, evidence of impairments in overt reflexive gaze following was evident in a well-matched comparison between children with and without autism, but not in a less well matched comparison. The absence of a validity effect among children with autism is consistent with post-hoc analyses following a marginally significant interaction by Johnson et al. (2005) but inconsistent with findings by Chawarska et al. (2003). Because a validity effect was apparent among typically developing children (only the older CA matches) but not those with
autism at the longer SOA after controlling for NVMA, the longer SOA may have captured more volitional aspects of gaze (see Ristic et al., 2002). Children with autism were faster to orient to targets regardless of cue validity at short SOAs but slower at longer SOAs.

**Hypothesis 2: impaired integration of gaze and affect.** Children with autism for whom RT data were available for all emotion validity pairings did not differ reliably from controls in the number of usable trials provided at either SOA. CA matched controls who supplied data for emotion integration analyses were higher in NVMA at the long SOA ($p = .048$). No reliable differences in CA or NVMA were observed at either SOA among NVMA matches (short $N = 16$; long $N = 15$) and children with autism (short $N = 11$; long $N = 13$). Analyses focus on NVMA controls as they are better matched.

A repeated measures analysis at the short SOA with validity and emotion (fear, happy or neutral) as factors revealed an interaction between diagnosis and validity $F (1, 25) = 4.376, p = .047$, a main effect of emotion $F (2, 50) = 5.279, p = .008$, and an interaction between emotion and diagnosis $F (2, 50) = 4.515, p = .016$. However, the predicted interaction between validity, emotion and diagnosis was not observed. Post-hoc tests revealed that participants oriented to targets more slowly following fearful ($M = 2.579, SE = .019$) than neutral gaze cues ($M = 2.529, SE = .017, p = .008$). Post-hoc tests investigating the relation between emotion and diagnosis were not conclusive but visual inspection of the data suggested that individuals with autism oriented faster than typically developing participants to neutral and fearful cues but oriented slower than typically developing participants to happy cues, regardless of cue validity.

A similar analysis of data from the longer SOA revealed a significant effect of emotion $F (2, 42) = 6.963, p = .002$ and an interaction between emotion and diagnosis $F (2, 52) = 3.601, p =$
Post-hoc tests revealed faster orienting to happy ($M=2.496$, $SE=.014$) relative to neutral ($M=2.533$, $SE=.013$; $p=.001$) or fearful ($M=2.536$, $SD=.016$; $p=.002$) gaze cues, regardless of cue validity. A significant main effect of emotion on usable trials was observed wherein fearful cues ($M=2.51$, $SD=.17$) yielded fewer usable trials than neutral cues ($M=2.98$, $SD=.153$; $F(2,76)=5.277$, $p=.007$). Post-hoc tests investigating the relation between emotion and diagnosis were again not conclusive but visual inspection of the data suggested that individuals with autism oriented faster than typically developing participants to fearful and happy cues but oriented slower than typically developing participants to neutral cues, regardless of cue validity.

Again contrary to our hypotheses, there was little support for the possibility that integration of gaze and emotion is a core deficit underlying gaze following difficulties in autism. At the same time, we observed no evidence of integration of emotion and gaze among typically developing children.

**Hypothesis 3: difficulty with the referential significance of gaze is related to developmental level rather than autism.**

**Gaze following.** Children with autism ($N=16$) did not differ reliably in CA or NVMA from CA matched controls ($N=18$). NVMA matched controls ($N=18$) were reliably lower in CA ($p=.003$). To compare performance of the ASC group to CA matches, we conducted a mixed ANOVA on first look difference scores (gaze following) with location of word learning as a repeated measure and diagnosis as a between subjects measure. Children with ASCs ($M=3.667$, $SE=.466$) followed gaze less frequently overall than typically developing children across contexts ($M=5.417$, $SE=.425$; $F(1,31)=7.694$, $p=.009$; see Figure 5). Gaze following occurred less frequently in the eye-tracker ($M=4.050$, $SE=.331$) than in person ($M=5.033$, $SE= .034$.
A similar ANCOVA conducted with NVMA matches (except that CA was entered as a covariate) revealed significant main effects of diagnosis \( F(1, 31) = 6.725, p = .014 \) and age \( F(1, 31) = 6.252, p = .018 \). Again, children with ASCs (\( M=3.172, SE= .479 \)) followed gaze less frequently overall than typically developing children (\( M=4.986, SE= .448; p = .014 \)).

As hypothesized, children with ASCs exhibited less gaze following across contexts than typically developing children, even after controlling for differences in developmental level.

**Eye-tracking word learning.** For CA and NVMA matches, eye-tracking referent preference (relative duration the referent was attended to during test) was unrelated to referent selection after eye-tracking (whether or not the child selected both referents correctly). Binary logistic regressions with CA or NVMA matches revealed no statistically significant associations between referent selection during test and diagnosis, first look difference scores (gaze following during training), age or NVMA. An ANCOVA conducted with CA matches predicting eye-
tracking referent preference from the same set of predictors revealed an association between first look difference scores $F(1, 28) = 8.770$, $p = .006$, NVMA $F(1, 28) = 6.371$, $p = .018$ and referent preference. An identical ANOVA with NVMA matches yielded no significant results. Thus, eye-tracking measures of word learning were not related to one another, referent selection after eye-tracking was not attributable to developmental level or autism, and referent preference during the eye-tracking test phase was associated with gaze following during training and cognitive level (at least for CA matches).

**In-person word learning.** For CA matches, NVMA was associated with in-person referent selection (whether or not the child selected the correct object on all test trials) ($B = 1.138$, $SE = .061; p = .034$). In-person first look difference scores, age, and diagnosis were not significantly associated with in-person referent selection. For NVMA matches, NVMA was associated with referent selection ($B = 1.124$, $SE = .055$, $p = .036$), but CA, diagnosis, and difference scores were not.

The third hypothesis, therefore was partially confirmed: gaze following difficulties were associated with autism while referent selection after in-person training and referent preference during eye-tracking were related to cognitive level. No relations between gaze behaviors during training and referent selection were observed although gaze following was associated with referent preference during eye-tracking.

**Discussion**

The current study examined three potential mechanisms underlying atypical gaze following in autism by using a range of measures to assess different aspects of gaze following in young children with and without autism. Results suggest that both high- and low-level
atypicalities of gaze following may be a core characteristic of autism in early childhood. In contrast, referential use of gaze may be more associated with developmental level than autism. No evidence for integration of gaze and affect was observed regardless of diagnosis.

**Impaired Low-level Gaze Following in Autism**

Reflexive gaze following typically begins to emerge within days of birth (Farroni et al., 2004), but the children with autism we observed (at a mean age of 4.5) exhibited less evidence of overt reflexive orienting to gaze than typically developing controls. Contrary to our hypotheses, only typically developing participants showed evidence of gaze cueing at a short SOA. These findings are consistent with statistically weaker effects observed with a smaller sample and a longer SOA by Johnson et al. (2005). Like Chawarska et al. (2010), we did not find an interaction between validity and diagnosis with CA matches. However, when our participants with and without ASCs were matched in cognitive level (and not significantly different in age), we found evidence of atypical reflexive gaze following in autism. Our sample size was larger and stringently matched but also older than the children assessed by Chawarska et al. (2010). It is possible that overt reflexive gaze following becomes more atypical with development in autism due to atypical social interactions, but this is somewhat unlikely given the early emergence of reflexive gaze following in typical development, Longitudinal assessment of reflexive gaze following among the infant siblings of children with autism might reveal whether atypicalities of social attention and gaze following emerge simultaneously or if one precedes the other.

Reduced gaze cueing at the longer SOA was apparent only when NVMA was controlled. Gaze cueing at longer SOAs may have a more volitional aspect than gaze following at short SOAs, which may be more reflexive (see Ristic et al., 2002).
Children with autism were faster to disengage from neutral faces regardless of cue validity at a short SOA but slower at a longer SOA. Atypically fast disengagement by children with autism was observed by Chawarska et al. (2010) at a short SOA and by Johnson et al. (2005) at a 1500 ms SOA. Variation in the speed at which children with and without autism disengage from faces at different SOAs calls into question the idea that children with autism may disengage more quickly because they are less interested in faces (e.g. Chawarska et al., 2010).

No Evidence of Impaired Integration of Gaze and Emotion

The data presented here did not suggest that children with autism have particular difficulty integrating gaze and affect. However, typically developing children also did not show evidence of gaze emotion integration. Previous research demonstrating gaze emotion integration was conducted with adults using covert attention cueing paradigms (e.g. Bayliss et al., 2010; Friesan et al., 2010; Graham et al., 2010; Heitanen & Leppänen, 2003). Enhanced reflexive orienting to emotional cues may emerge later in development. Because it is not always observed even in adulthood, some have asserted that emotion and gaze are processed independently early in visual processing (Heitanen & Leppänen, 2003). Others have found evidence that the integration of gaze cues and emotion does occur, particularly when targets are emotionally salient (Bayliss et al., 2010; Friesan et al., 2010; Pecchinenda et al., 2008). However, emotionally enhanced validity effects have also been observed with neutral targets (Graham et al., 2010; Tipples et al., 2010). Future research should examine gaze and emotion integration across typical development using affectively charged and neutral targets to develop normative standards to compare with atypical development. Relating performance on eye-tracking measures of emotion integration to expressive emotion integration (e.g. Kasari et al., 1990) would elucidate if similar or different mechanisms underlie the two.
Gaze Following but not Reference Associated with Autism

Contrary to what one would expect if reduced understanding of the referential significance of gaze underlies gaze following deficits in autism, reduced gaze following was apparent among children with autism while word learning was related to cognitive level rather than autism. Relations between cognitive level and word learning were only apparent during a relatively difficult word learning task wherein object location varied during test and not when object location did not vary during test (i.e. referent selection after eye-tracking). Interestingly, referent selection was not associated with gaze following, while preferentially looking toward the referent when it was labeled was associated with both gaze following and cognitive level. The current findings confirm the Gliga et al. (in press) claim that gaze following is “necessary but not sufficient for successful word learning” by demonstrating that frequency of gaze following may be less related to word learning in response to gaze cues (at least as indexed by compliance with requests to indicate a labeled object) than cognitive level.

Limitations

The majority of the participants in the current study were high functioning (only 4 had a NVIQ at or below 70). They were thus comparable to children assessed by Luyster and Lord (2009). The current study suggests that subtle difficulties learning words from gaze are attributable to developmental level. Comparisons of a range of children with autism (including more with severe intellectual delays) to typically developing and developmentally delayed children would yield further insights into links between word learning from social cues and developmental level in autism. Additional assessments of memory for words learned in this manner could be administered after a delay to clarify relations between gaze following, diagnosis
and word retrieval rather than word recognition.

Infrequent associations between gaze following and word learning in this study may be attributable to overly easy word learning tasks. While other studies with similar populations trained similar numbers of objects and used more clues to reference, such as carrier phrases or gestures (i.e. Luyster & Lord, 2009; Parish-Morris et al., 2007), a greater variety of training objects (such as used by Akechi et al., 2011) might yield more sensitive measures of word learning. The lack of correspondence between the eye-tracking measure of word learning based on looking behaviors and the measure based on referent selection implies that care should be taken to validate eye-tracking measures of gaze following and word learning.

**Conclusion**

The current study provides evidence that the development of gaze following may be atypical rather than simply delayed in autism: low-level gaze following that typically begins to emerge in infancy is not evident among young children with autism. Despite subtle difficulties with both high- and low-level aspects of gaze following among children with ASCs, word learning following gaze was associated with developmental level rather than autism. Thus, children with autism may “hack into” reference despite atypical gaze following. Interventions could focus on reflexive gaze following (perhaps using gaze contingent video games) and techniques for hacking into reference.

Given associations between joint attention and language development in autism and the fact that children diagnosed today achieve better language outcomes than in the past (Chakrabarti & Fombonne, 2001; Charman, Drew, Baird, & Baird, 2003), it would be intriguing to conduct a large-scale assessment of overt reflexive gaze following in people with autism at a broad range
of ages to determine if gaze following atypicalities are also less apparent among younger relative to older individuals with autism.

The current study demonstrates the importance of controlling developmental level when comparing different paradigms believed to measure higher-and lower-level aspects of core symptoms in autism. This developmental approach revealed surprising atypicalities in low-level gaze following, but not emotion gaze integration or an understanding of reference, among young children with autism.
References


99


* Shared first authorship


